

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE



In re: S. Patent Application of

YASUDA et al.

Application Number: 10/767,778

Filed: January 30, 2004

For: VIDEO CAMERA

ATTORNEY DOCKET NO. NITT.0180

Honorable Assistant Commissioner  
for Patents  
Washington, D.C. 20231

LETTER

Sir:

The below-identified communications are submitted in the above-captioned application or proceeding:

- |                                     |                           |   |
|-------------------------------------|---------------------------|---|
| <input checked="" type="checkbox"/> | Priority Documents (1)    |   |
| <input checked="" type="checkbox"/> | Request for Priority      | <input type="checkbox"/> Assignment Document              |
| <input type="checkbox"/>            | Response to Missing Parts | <input type="checkbox"/> Petition under 37 C.F.R. 1.47(a) |
|                                     | w/ signed Declaration     | <input type="checkbox"/> Check for \$130.00               |

- ☒ The Commissioner is hereby authorized to charge payment of any fees associated with this communication, including fees under 37 C.F.R. § 1.16 and 1.17 or credit any overpayment to Deposit Account Number 08-1480. A duplicate copy of this sheet is attached.

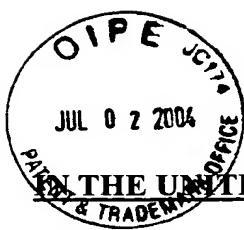
Respectfully submitted,

Stanley P. Fisher  
Registration Number 24,344

---

Juan Carlos A. Marquez  
Registration Number 34,072

REED SMITH LLP  
3110 Fairview Park Drive  
Suite 1400  
Falls Church, Virginia 22042  
(703) 641-4200  
July 2, 2004



**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re U.S. Patent Application of )  
YASUDA et al. )  
Application Number: 10/767,778 )  
Filed: January 30, 2004 )  
For: VIDEO CAMERA )  
ATTORNEY DOCKET NO. NITT.0180 )

Honorable Assistant Commissioner  
for Patents  
Washington, D.C. 20231

**REQUEST FOR PRIORITY  
UNDER 35 U.S.C. § 119  
AND THE INTERNATIONAL CONVENTION**

Sir:

In the matter of the above-captioned application for a United States patent, notice is hereby given that the Applicant claims the priority date of January 30, 2003, the filing date of the corresponding Japanese patent application 2003-021301.

A certified copy of Japanese patent application 2003-021301 is being submitted herewith. Acknowledgment of receipt of the certified copy is respectfully requested in due course.

Respectfully submitted,

Stanley P. Fisher  
Registration Number 24,344

**REED SMITH LLP**  
3110 Fairview Park Drive  
Suite 1400  
Falls Church, Virginia 22042  
(703) 641-4200  
July 2, 2004

Juan Carlos A. Marquez  
Registration Number 34,072

**PATENT OFFICE**  
**JAPANESE GOVERNMENT**

This is to certify that the annexed is a true copy of the following application as filed with this office.

Date of Application : January 30, 2003  
Application Number : Patent Application No. 2003-021301  
Applicant (s) : Hitachi, Ltd.

Dated this 23rd day of January, 2004

---

Yasuo IMAI  
Commissioner,  
Patent Office

Certificate No. 2004-3002029



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日            2 0 0 3 年    1 月 3 0 日  
Date of Application:

出 願 番 号            特 願 2 0 0 3 - 0 2 1 3 0 1  
Application Number:  
[ST. 10/C] :            [ J P 2 0 0 3 - 0 2 1 3 0 1 ]

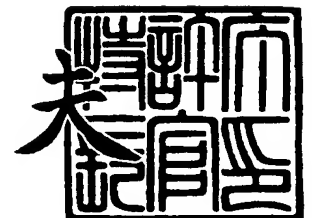
出    願    人            株 式 会 社 日 立 製 作 所  
Applicant(s):



2 0 0 4 年    1 月 2 3 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康



【書類名】 特許願

【整理番号】 H02018631A

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 12/00

【発明者】

    【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

    【氏名】 保田 淑子

【発明者】

    【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

    【氏名】 川本 真一

【発明者】

    【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

    【氏名】 江端 淳

【発明者】

    【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所中央研究所内

    【氏名】 沖津 潤

【特許出願人】

    【識別番号】 000005108

    【氏名又は名称】 株式会社 日立製作所

【代理人】

    【識別番号】 100075096

    【弁理士】

    【氏名又は名称】 作田 康夫

    【電話番号】 03-3212-1111

## 【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 分散ファイルシステムのファイルレプリケーション方法及び装置

【特許請求の範囲】

【請求項 1】

ネットワークに接続した複数のネットワークストレージ装置とレプリケーション装置とを含み、該レプリケーション装置はレプリケーション元のネットワークストレージ装置に存在する全ファイル及びディレクトリの属性情報を管理する管理テーブルを有する分散ファイルシステムにおいて、レプリケーション元のネットワークストレージ装置のデータの一部コピーをレプリケーション先のネットワークストレージに作成するファイルレプリケーション方法であって、

予め前記レプリケーション装置内にレプリケーション対象のファイルを特定するレプリケーション情報を記録するステップと、

クライアントからのファイルアクセス要求を受け付けるステップと、

前記管理テーブルと前記レプリケーション情報を使用して、前記ファイルアクセス要求の実行に伴ってレプリケーション動作を行なうか否かを判定するステップと、

前記判定の結果がレプリケーション動作を行なうという場合には、該ファイルアクセス要求を前記レプリケーション元のネットワークストレージ装置と、前記レプリケーション先のネットワークストレージ装置に対して同時に転送するステップとを有するファイルレプリケーション方法。

【請求項 2】

前記レプリケーション装置は、レプリケーション元ネットワークストレージ装置と、レプリケーション先ネットワークストレージ装置の間で、コピー対象のファイル及びディレクトリの内容が一致するか否かを示す同期情報を含み、前記判定ステップでは前記同期情報が一致を示すこと一つの条件にレプリケーション動作を行うと判定することを特徴とする請求項 1 記載のファイルレプリケーション方法。

【請求項 3】

前記判定ステップでは、受け付けたファイルアクセス要求が書き込み要求であることを一つの条件にレプリケーション動作を行なうと判定することを特徴とする請求項1記載の分散ファイルシステムのファイルレプリケーション方法。

#### 【請求項4】

ネットワークに接続した複数のネットワークストレージ装置の間でファイルレプリケーションを行なうレプリケーション装置であって、

レプリケーション元のネットワークストレージ装置に存在する全ファイル及びディレクトリの属性情報を管理する管理テーブルと、

レプリケーションするかどうかを示すレプリケーション情報を保持する手段と

、  
クライアントからのファイルアクセス要求を一元的に受け付ける要求受付手段と、

該ファイルアクセス要求を複数のネットワークストレージ装置に同時に転送するレプリケーション手段を備え、

該レプリケーション手段は、

該管理テーブルとレプリケーション情報を利用してレプリケーションするかどうかを判定する判定手段と、

判定結果に従い、該ファイルアクセス要求を、レプリケーション元のネットワークストレージ装置と、レプリケーション先のネットワークストレージ装置に対して同時に転送する要求処理手段を含むことを特徴とするレプリケーション装置

。

#### 【請求項5】

前記レプリケーション手段は、さらにレプリケーション元のネットワークストレージ装置とレプリケーション先のネットワークストレージ装置の間で、コピー対象の全ファイルおよび全ディレクトリの整合性をとる整合手段を含むことを特徴とする請求項4記載のレプリケーション装置。

#### 【請求項6】

前記判定手段は、受け付けたファイルアクセス要求が書き込み要求である場合にレプリケーションを行なうと判定することを特徴とする請求項4記載のレプリ



ケーション装置。

**【請求項 7】**

前記レプリケーション装置は、前記レプリケーション元のネットワークストレージ装置と前記レプリケーション先のネットワークストレージ装置間でコピー対象のファイルおよびディレクトリの内容が一致することを示す同期情報を保持する手段を更に有し、前記判定手段は、該同期情報も使用してレプリケーションするかどうかを判定することを特徴とする請求項 4 記載のレプリケーション装置。

**【請求項 8】**

前記レプリケーション情報は、特定ユーザ識別子あるいはグループ識別子を持つファイル、あるいは特定ディレクトリに所属するファイル、あるいは特定ファイル識別子を持つファイルがコピー対象であることを示す少なくとも一つ以上のルールであることを特徴とする請求項 4 記載のレプリケーション装置。

**【請求項 9】**

複数のネットワークストレージと、これらネットワークストレージに分散して存在するファイル及びディレクトリの構造と属性情報を一元管理し、これらネットワークストレージを外部から一元的にアクセス可能とする仮想一元化装置とを含む仮想一元化ファイルシステムのデータの一部コピーを外部ネットワークストレージに作成するファイルレプリケーション方法であって、

予め前記仮想一元化装置内にレプリケーション対象のファイルを特定するレプリケーション情報を記録するステップと、

クライアントから前記仮想一元化ファイルシステムへのファイルアクセス要求を一元的に受け付けるステップと、

ファイルごとにネットワークストレージを決定して設定するマッピング手段により前記ファイルアクセス要求の対象ファイルの格納先のネットワークストレージを特定するステップと、

前記レプリケーション情報を用い、前記ファイルアクセス要求の実行に伴ってレプリケーション動作を行なうか否かを判定するステップと、

特定したネットワークストレージに前記ファイルアクセス要求を転送するとともに、前記判定の結果がレプリケーション動作を行なうという場合には、該ファ

イルアクセス要求をレプリケーション先の外部ネットワークストレージにも転送し、もって前記対象ファイルの格納先のネットワークストレージとレプリケーション先の外部ネットワークストレージの双方で該ファイルアクセス要求が要求するファイルアクセスを実行させるステップとを有するファイルレプリケーション方法。

【請求項 10】

前記判定ステップでは、受け付けたファイルアクセス要求が書き込み要求であることを一つの条件にレプリケーション動作を行なうと判定することを特徴とする請求項 9 記載のファイルレプリケーション方法。

【請求項 11】

前記ファイル格納先のネットワークストレージ装置からの該ファイルアクセス要求に対する応答と、前記レプリケーション先のネットワークストレージからの該ファイルアクセス要求に対する応答を収集し1つにまとめて前記クライアントに返送するステップを更に含むことを特徴とする請求項 9 記載のファイルレプリケーション方法。

【請求項 12】

ネットワークに接続された複数のネットワークストレージを仮想的に単一のファイルシステムに見せる複数ネットワークストレージの仮想一元化装置であって、

前記仮想一元化ファイルシステムに存在する全ファイル及びディレクトリの構造と属性情報を管理する一元管理ディレクトリと、

外部ファイルシステムをレプリケーション先としてレプリケーションすべきファイルを特定するレプリケーション情報を保持する手段と、

クライアントからのファイルアクセス要求を受け付ける要求受付手段と、

前記ファイルアクセス要求のファイルアクセス対象のネットワークストレージを決定するマッピング手段と、

該一元管理ディレクトリとレプリケーション情報を使用して前記ファイルアクセス要求のファイルアクセス対象を前記外部ファイルシステムにレプリケーションするかどうかを判定する判定手段と、

判定手段によりレプリケーションを行なうと判定した場合、前記ファイルアクセス要求を、レプリケーション先の外部ファイルシステムと、該マッピング手段により決定したネットワークストレージに対して同時に、前記ファイルアクセス要求を転送する要求転送手段とを有することを特徴とする複数ネットワークストレージの仮想一元化装置。

【請求項 1 3】

前記仮想一元化ファイルシステムと前記レプリケーション先の外部ファイルシステムの間で、コピー対象のファイルおよびディレクトリの整合性をとる整合手段を更に有することを特徴とする請求項 1 2 記載の複数ネットワークストレージの仮想一元化装置。

【請求項 1 4】

前記仮想一元化ファイルシステムに属するネットワークストレージ装置からのファイルアクセス要求に対する応答と、前記レプリケーション先の外部ファイルシステムのからのファイルアクセス要求に対する応答をあつめて1つの応答にしてクライアントに返送する応答収集手段を更に有することを特徴とする請求項 1 2 記載の複数ネットワークストレージの仮想一元化装置。

【請求項 1 5】

前記仮想一元化ファイルシステムと前記レプリケーション先のファイルシステム間で、コピー対象の全ファイルおよびディレクトリが一致することを示す同期情報を保持する手段を更に有し、前記判定手段が、該同期情報も使用してレプリケーションするかどうかを判定することを特徴とする請求項 1 2 記載の複数ネットワークストレージの仮想一元化装置。

【請求項 1 6】

前記判定手段は、ファイルアクセス要求が読み出し要求である場合にはレプリケーションしないと判定し、レプリケーション先の外部ファイルシステムには該ファイルアクセス要求を転送しないことを特徴とする請求項 1 2 記載の複数ネットワークストレージの仮想一元化装置。

【請求項 1 7】

前記レプリケーション情報は、特定ユーザ識別子あるいはグループ識別子を持

つファイル、あるいは特定ディレクトリ配下のファイル、あるいは特定ファイル識別子を持つファイルがコピー対象であること示す少なくとも一つ以上のルールである請求項 12 記載の複数ネットワークストレージの仮想一元化装置。

**【請求項 18】**

一元管理ディレクトリで管理されるファイルおよびディレクトリがマスタであることを示すマスタ情報を保持する手段を更に有し、前記判定手段が、該マスタ情報にも従ってレプリケーションするかどうかを判定することを特徴とする請求項 12 記載の複数ネットワークストレージの仮想一元化装置。

**【請求項 19】**

前記レプリケーション情報は、該ルールに加えて、該ルールを適用先の仮想一元化ファイルシステムを識別する情報も含むことを特徴とする請求項 17 記載の複数ネットワークストレージの仮想一元化装置。

**【請求項 20】**

複数のネットワークストレージと仮想一元化装置から構成する仮想一元化ファイルシステムにおいて、

該仮想一元化装置は、レプリケーション手段を有し、

第一の仮想一元化ファイルシステムは、クライアントからのアクセス要求に回答して、レプリケーション手段を起動し、該第一の仮想一元化ファイルシステムの部分コピーをネットワークに接続した異なる仮想一元化ファイルシステムにリアルタイムに作成することを特徴とする仮想一元化ファイルシステム。

**【請求項 21】**

ネットワークに接続し、複数のネットワークストレージと仮想一元化装置から構成する仮想一元化ファイルシステムにおいて、

該仮想一元化装置は、レプリケーション手段を有し、

第一の仮想一元化ファイルシステムは、クライアントからのアクセス要求に回答して、レプリケーション手段を起動し、該第一の仮想一元化ファイルシステムの部分コピーを、ネットワークに接続した別のネットワークストレージにリアルタイムに作成することを特徴とする仮想一元化ファイルシステム。

**【発明の詳細な説明】**

## 【0001】

## 【発明の属する技術分野】

本発明は、複数ネットワークストレージ装置間のファイルレプリケーション方法およびファイルレプリケーション機能を備えたネットワークストレージ装置に関する。

## 【0002】

## 【従来の技術】

従来、データはクライアント計算機に直接接続したストレージ (DAS: Direct Attached Storage) に格納されていた。DAS上のデータは直接接続しているクライアント計算機を介してのみアクセス可能であるため、クライアント計算機が他のクライアント計算機に接続したDASのデータをアクセスする場合には、DASが接続している計算機を経由してデータをアクセスする必要があった。

## 【0003】

しかしながら、近年のネットワーク技術の発展により個々のクライアント計算機に格納すべきデータ量が飛躍的に増大している。それに伴い、これらのデータを効率良く管理したいというニーズも増大し、DASに代わりネットワークに接続したストレージシステム (ネットワークストレージ) が導入されるようになってきた。ネットワークストレージを導入することにより、複数のクライアント計算機間でデータを共有できるため、管理者は共有データを効率良く管理することができる。

## 【0004】

ネットワークストレージの例としては、SAN (Storage Area Network) で接続されてブロックアクセスを提供するSANストレージ、IPネットワークやInfiniband等で接続されてファイルアクセスを提供するNAS (Network Attached Storage) がある。その中でもNASは、SANストレージに比べて安価であり、かつNFS (Network File System) やCIFS (Common Internet File System) のような複数の標準的なファイルアクセスプ

ロトコルを用いて複数のクライアント計算機が簡単にファイルを共有することができるため、その市場は拡大している。

#### 【0005】

このように、ネットワークストレージは、管理者が効率的に複数クライアント計算機の共有データを一括して管理できるため、世の中に広く受け入れられている。一方、ネットワークストレージに格納された共有データは、管理者によりテープドライブ等を使用してバックアップされ、必要に応じてテープドライブから復旧していた。しかしながら、テープドライブは低速であるため、それを利用したバックアップや復旧は非常に時間がかかり、この時間は共有データ量に比例して増大してしまう。特に、近年のハードディスク技術の進展により、ネットワークストレージの容量もますます増大しており、低速なテープバックアップは効率が悪い。

#### 【0006】

低速なテープバックアップの問題点を解決するため、近年、ネットワークストレージに保持した共有データの複製をテープドライブではなくネットワークに接続した別のネットワークストレージに保持する方法が用いられてきている。この方法により、テープドライブを利用したバックアップよりも高速に、共有データの複製を作成することができ、かつ高速に復旧できる。

#### 【0007】

例えば、非特許文献1には、複数のネットワークストレージにおいて、リアルタイムにファイルのコピーを作成する方法が開示されている。VERITAS File Replicatorは、クライアントからの更新系のファイルアクセス要求を受け取ると、レプリケーションの対象となる第一のファイルシステムとレプリケーション先の複数のファイルシステムに同時にその要求を転送する。レプリケーションは、対象となるファイルシステム全体に対して行なう。

また、特許文献1には、ジャーナリングファイルシステムと呼ぶファイルレプリケーションのための特別な機能を持つ仮想一元化スイッチが、プライマリの仮想一元化ファイルシステムを構成するサーバ群と、セカンダリのファイルサーバの両方に対して、更新系のファイルアクセス要求を転送してファイルシステム全

体のコピーを生成する方法が開示されている。ジャーナリングファイルスイッチは、プライマリの仮想一元化ファイルシステムを構成するサーバ群から応答を受け取ると、セカンダリのファイルサーバにおいて更新系リクエストが処理されるのを待たずに、その応答をクライアントに返送する。それにより、ファイルコピーのオーバーヘッドを削減することができる。

さらに、特許文献 2 には、分散型のコンピュータシステムにおいて、レプリケーション対象のコンピュータシステムのサブツリーのみをコピーするレプリケーション装置が開示されている。このレプリケーション装置では、クライアントからのファイルアクセス要求は、ファイルを持つコンピュータシステムに転送される。コンピュータシステムは、更新系のファイルアクセス要求である場合、変更があったことをレプリケートド記憶変更記録簿と呼ぶログに登録する。レプリケートド記憶変更記録簿に登録された更新系リクエストは、一定の間隔あるいはあらかじめスケジュールされた間隔で、レプリケーション先のコンピュータシステムにコピーされる。

#### 【非特許文献 1】

VERITAS File Replicator 3.0.3, System Administrator's Guide, 2 0 0 1 年 6 月、VERITAS Software Corporation 発行

#### 【特許文献 1】

米国特許公開 2 0 0 2 / 0 1 2 0 7 6 3 号

#### 【特許文献 2】

特開平 0 7 - 2 1 9 8 3 0 号公報

#### 【0 0 0 8】

#### 【発明が解決しようとする課題】

上記非特許文献 1 及び特許文献 1 に記載された技術では、複数のネットワークストレージ間あるいは複数の仮想一元化ファイルシステム間でリアルタイムにファイルのコピーを作成することができる。しかしながら、レプリケーションはファイルシステム全体に対して行なう必要がある。

#### 【0 0 0 9】

先に述べたように、ネットワークストレージの容量はますます増大しているた

め、ファイルシステム全体のコピーしか作成できない場合には、バックアップ時間が増大してしまい問題である。さらに、共有データにはバックアップが必須な重要ファイルとバックアップの必要がないファイルが含まれる。しかしながら、上記第1および第2の公知例では、管理者が重要ファイルのみのバックアップを取得すればよい場合でも、ファイルシステム全体のコピーしか作成できないため、バックアップ時間がかかるだけでなく、コピー先のネットワークストレージの容量も十分に用意しなければならないことが問題である。

#### 【0010】

一方、上記特許文献2に記載されているファイルレプリケーション方法はファイルシステムのサブツリーのコピーを作成する方法を提供しているが、クライアントからのファイルアクセス要求に応答して、コピーをリアルタイムに作成することができない。そのため、障害が発生するような場合に最新データのバックアップが保持されていないことが問題である。

#### 【0011】

この問題点を解決するために、複数のネットワークストレージにおいて、ファイルシステムの部分コピーをリアルタイムに作成するファイルレプリケーション方法を提供する必要がある。

#### 【0012】

また、複数の仮想一元化ファイルシステムにおいても、仮想一元化ファイルシステムの部分コピーをリアルタイムに作成するファイルレプリケーション方法を提供する必要がある。

#### 【0013】

さらに、上記特許文献2では、選択的なファイルレプリケーション実行中の複数の仮想一元化ファイルシステムにおいて、仮想一元化ファイルシステムの一部に障害が発生したり、仮想一元化ファイルシステムが接続したネットワークに障害が発生し、障害から復旧するときに、管理者が手動で整合性をとる必要がある。手動で整合性をとるということは、管理者がバックアップをとったすべてのファイルについて、コピー済みか否かを調査する必要がある。

#### 【0014】



この問題点を解決するために、選択的なファイルレプリケーション実行中の複数の仮想一元化ファイルシステムにおいて、仮想一元化ファイルシステムあるいはネットワークに障害が発生し、障害復旧する場合に、複数の仮想一元化ファイルシステム間で自動でコピー対象のファイルおよびディレクトリの整合性をとって、レプリケーションを再開する方法を提供する必要がある。

#### 【0015】

本発明の第1の課題は、複数のネットワークストレージ間でレプリケーションの対象となるファイルシステムの部分コピーをリアルタイムに作成するファイルレプリケーション方法を提供することにある。

#### 【0016】

本発明の第2の課題は、複数の仮想一元化ファイルシステム間でレプリケーションの対象となる仮想一元化ファイルシステムの部分コピーをリアルタイムに作成するファイルレプリケーション方法を提供することにある。

#### 【0017】

本発明の第3の課題は、部分コピー実行中の複数のネットワークストレージあるいは複数の仮想一元化ファイルシステムにおいて、ファイルシステムの一部あるいはファイルシステムが接続するネットワークに障害が発生し、障害から復旧する場合でも、簡単にオリジナルのファイルおよびディレクトリとコピーの整合性をとる方法を提供することにある。

#### 【0018】

##### 【課題を解決するための手段】

本発明の第1の課題は、複数ネットワークストレージで構成された分散システムに、レプリケーション対象ファイルシステムのファイルディレクトリに関する属性情報を管理する管理テーブルと、コピーするかどうかの条件を示すレプリケーション情報と、クライアントからのファイルアクセス要求を受け付ける要求受付手段と、受け付けたファイルアクセス要求が更新系リクエストである場合に管理テーブルとレプリケーション情報を使用してレプリケーションするかどうかを判定する判定手段と、判定結果に従って、レプリケーション対象ネットワークストレージ装置のファイルシステムと、レプリケーション先ネットワークストレージ

ジのファイルシステムに対して、同時にそのファイルアクセス要求を転送する要求処理手段を設けることで解決できる。

#### 【0 0 1 9】

本発明の第2の課題は、レプリケーション元の仮想一元化ファイルシステムに属する仮想一元化装置に、レプリケーション元の仮想一元化ファイルシステムのファイルディレクトリツリー構造とファイル識別子を管理する一元管理ディレクトリと、ファイルをコピーするかどうかの条件を示すレプリケーション情報と、クライアントからのファイルアクセス要求を一元的に受け付ける要求受付手段と、ファイルアクセス要求の対象ファイルを保持するネットワークストレージを決定するマッピング手段と、受け付けたファイルアクセス要求が更新系リクエストである場合に、一元管理ディレクトリの管理する属性情報とレプリケーション情報を利用してレプリケーションするかどうかを判定する判定手段と、判定結果にしたがって、レプリケーション元の仮想一元化ファイルシステムに属し、マッピング手段により決定したネットワークストレージだけでなく、レプリケーション先の仮想一元化装置に対して、同時にファイルアクセス要求を転送する要求処理手段を設けることで解決できる。

#### 【0 0 2 0】

本発明の第3の課題は、第2の課題を解決する手段に加えて、レプリケーション対象の仮想一元化ファイルシステムに属する仮想一元化装置に、レプリケーション元のコピー対象のファイルおよびディレクトリと、レプリケーション先の仮想一元化ファイルシステムに保持するファイルおよびディレクトリとの整合性をとる整合手段を設けることで解決できる。

#### 【0 0 2 1】

複数のネットワークストレージで構成される分散システムにおいて、要求受付手段が一括してクライアントからのファイルアクセス要求を受け付け、管理テーブルを利用してレプリケーションするかどうかを判定し、コピー対象のファイルに対するファイルアクセス要求のみを、レプリケーション元ネットワークストレージと、レプリケーション先ネットワークストレージの両方に並行して転送することで、レプリケーション元ネットワークストレージのファイルシステムの部分

コピーをリアルタイムに別のネットワークストレージに作成することができる。

#### 【0 0 2 2】

また、複数ネットワークストレージと仮想一元化装置で構成される仮想一元化ファイルシステムにおいて、クライアントからファイルアクセス要求が発行されると、仮想一元化装置がそのファイルアクセス要求を受け付け、一元管理ディレクトリを利用して、アクセス対象ファイルの存在するネットワークストレージを特定し、かつアクセス対象のファイルがコピー対象であるかどうかを判定し、コピー対象のファイルに対するファイルアクセス要求である場合にのみ、アクセス対象のネットワークストレージとレプリケーション先仮想一元化ファイルシステムの仮想一元化装置に対してファイルアクセス要求を並行して転送することで、レプリケーション元仮想一元化ファイルシステムの部分コピーをリアルタイムに別の仮想一元化ファイルシステムに作成することができる。

#### 【0 0 2 3】

さらに、レプリケーション対象仮想一元化ファイルシステムを構成する仮想一元化装置が、コピー元のファイルと、レプリケーション先仮想一元化ファイルシステムに格納したファイルとの整合性をとる手段を持つため、障害発生によりレプリケーション元とレプリケーション先の仮想一元化ファイルシステム間でコピー対象のファイルおよびディレクトリの整合性が崩れても、簡単に再度整合性をとることができる。

#### 【0 0 2 4】

##### 【発明の実施の形態】

##### <代表的実施の形態>

図 1 に本発明の実施の形態であるネットワークストレージシステム全体の構成例を示す。ネットワークストレージシステムは、クライアント 1 0 0 ～ 1 0 1、ネットワーク 1 0 2、レプリケーション元仮想一元化ファイルシステム 1 0 3、レプリケーション先仮想一元化ファイルシステム 1 0 4 で構成する。本実施の形態では、クライアント 1 0 0 ～ 1 0 1 がレプリケーション元仮想一元化ファイルシステム 1 0 3 を使用してファイル共有を行なっている場合に、レプリケーション元仮想一元化ファイルシステム 1 0 3 のファイルおよびディレクトリの複製を

レプリケーション先仮想一元化ファイルシステム 104 に生成する方法について詳細に説明する。レプリケーション元仮想一元化ファイルシステム 103 とレプリケーション先仮想一元化ファイルシステム 104 の構成は同じであるが、レプリケーション先仮想一元化ファイルシステム 104 はクライアント 100～101 がアクセスしないように設定する。具体的には、`exports` ファイルにレプリケーション元仮想一元化ファイルシステム 103 を登録しない。また、レプリケーション先仮想一元化ファイルシステム 104 では、レプリケーションに関する設定情報の値が異なる。設定情報については後述する。また、レプリケーション元仮想一元化ファイルシステム 103 とレプリケーション先仮想一元化ファイルシステム 104 の全体容量は同じである必要はない。レプリケーション先仮想一元化ファイルシステム 104 は、レプリケーション元仮想一元化ファイルシステム 103 のコピー対象のディレクトリおよびファイルを保持できるだけの容量をもてばよい。

#### 【0025】

本実施の形態では、レプリケーション先はレプリケーション元と同様に仮想一元化ファイルシステムであるが、コピー対象のディレクトリおよびファイルを保持できさえすれば、レプリケーション先は大容量のネットワークストレージでもよい。

#### 【0026】

クライアント 100～101 は、NFS (Network File System) クライアントプログラムあるいは、CIFS (Common Internet File System) を介して、ネットワーク 102 経由で、レプリケーション元仮想一元化ファイルシステム 103 にアクセスする。ここで、クライアントとして NFS と CIFS のみを例にとったが、これ以外の標準的なファイルアクセスプロトコルを利用してもよい。その他のファイルアクセスプロトコルとしては、AFP (Apple File Protocol) や NCP (Netware Core Protocols) 等がある。

ネットワーク 102 は、IP ネットワークであれば、ローカルエリアネットワーク (LAN) でも、ワイドエリアネットワーク (WAN) でもよい。

**【 0 0 2 7 】**

レプリケーション元仮想一元化ファイルシステム 1 0 3 は、仮想一元化装置 1 1 0、内部ネットワーク 1 1 5、ネットワークストレージ 1 2 0～1 2 2 で構成する。

**【 0 0 2 8 】**

ネットワークストレージ 1 2 0～1 2 2 は、専用装置であってもよいし、ストレージ装置を有する汎用サーバ、ワークステーション、あるいは P C にリモート制御手段を搭載してもよい。本実施の形態では、ネットワークストレージ 1 2 0～1 2 2 は内部ネットワーク 1 1 5 に接続しているが、ネットワーク 1 0 2 に接続してもよい。

**【 0 0 2 9 】**

ネットワークストレージ 1 2 0 は、リモート制御手段 1 3 0、ストレージ装置 1 4 0 で構成する。ネットワークストレージ 1 2 1～1 2 2 の構成もネットワークストレージ 1 2 0 と同様である。

**【 0 0 3 0 】**

リモート制御手段 1 3 0 は、内部ネットワーク 1 1 5 に接続し、クライアント 1 0 0～1 0 1 からのマウント要求およびファイルの生成要求、読み出し要求、書きこみ要求、あるいはディレクトリの生成、ディレクトリの削除等を行なう。マウントとは、ネットワークを介してアクセスするネットワークストレージのファイルシステムをクライアント計算機のファイルシステムの一部にする処理である。マウント処理により、クライアントはネットワーク上に分散するストレージ装置をあたかもクライアントのローカルファイルシステムのようにアクセスできる。リモート制御手段 1 3 0 は、N F S あるいは C I F S 等の標準的なファイルアクセスプロトコルを使用して、クライアント 1 0 0～1 0 1 と通信を行なう。例えば、N F S プロトコルを使用する場合には、公知の `mountd` と `nfsd` を使用する。

**【 0 0 3 1 】**

クライアント 1 0 0～1 0 1 は、ストレージ装置に対してファイルハンドルを利用してファイルアクセス要求を発行する。ファイルハンドルはファイルオブジ

エクトの識別子であり、全ファイルに対して一意であることが保証されている。ファイルオブジェクトとはファイルあるいはディレクトリのいずれかである。ファイルハンドルは、親ディレクトリの i ノード番号とファイルオブジェクトの i ノード番号を含む。

#### 【 0 0 3 2 】

N F S プロトコルで規定されたファイルやディレクトリに対するアクセス要求としては、ファイル名を与えてファイルハンドルを取得する L O O K U P 要求、ファイルを作成する C R E A T E 要求、ファイルを読み出す R E A D 要求、ファイルを書きこむ W R I T E 要求、ファイルの属性を設定する S E T A T T R 要求、ファイルの属性を読み出す G E T A T T R 要求、ファイルを削除する R E M O V E 要求、ファイル名を変更する R E N A M E 要求、ディレクトリを作成する M K D I R 要求、ディレクトリを読み出す R E A D D I R 要求、およびディレクトリを削除する R M D I R 要求がある。このうち更新系のファイルアクセス要求は、C R E A T E 要求、W R I T E 要求、S E T A T T R 要求、R E M O V E 要求、R E N A M E 要求、M K D I R 要求、R M D I R 要求である。そのほかにも、更新系のファイルアクセス要求としては、L I N K 要求、S Y M L I N K 要求がある。

#### 【 0 0 3 3 】

リモート制御手段 1 3 0 は、マウント要求および L O O K U P 要求処理の結果得られるファイルハンドルを引数として、クライアントからのファイルアクセス要求を受け付け、各要求を処理して結果をクライアントに返す。

#### 【 0 0 3 4 】

C I F S プロトコルを使用する場合には、公知の S a m b a 等を使用すればよい。本実施の形態ではクライアント 1 0 0 ~ 1 0 1 が N F S プロトコルを利用してレプリケーション元仮想一元化ファイルシステム 1 0 3 にアクセスするものとする。

#### 【 0 0 3 5 】

ストレージ装置 1 4 0 にはファイルシステムを搭載する。U N I X (登録商標) 系のオペレーティングシステムとしては、サン・マイクロシステムズ社に S o

laris、インターナショナル・ビジネス・マシーン社 (International Business Machines Corporation) 社の AIX、ヒューレット・パカード社 (Hewlett-Packard Company) の HP-UX、Linux、及び、FreeBSD 等がある。

#### 【0036】

仮想一元化装置 110 は、ネットワークストレージ 120～122 を仮想的に一元管理し、クライアント 100～101 からは仮想的に一つのファイルシステムに見えるようにする。すなわち、クライアント 100～101 に対して、単一のファイルディレクトリツリー構造を見せる。本実施の形態では、管理者があらかじめコピー対象として設定するファイルに対するファイルアクセス要求を仮想一元化装置 110 に送信すると、仮想一元化装置 110 がそのアクセス要求を、リアルタイムにレプリケーション元仮想一元化ファイルシステム 103 のいずれかのネットワークストレージと、レプリケーション先仮想一元化ファイルシステム 104 に対して並行して転送することに特徴がある。一方、ディレクトリに対するファイルアクセス要求の場合、仮想一元化装置 110 は、そのアクセス要求をリアルタイムにレプリケーション元仮想一元化ファイルシステム 103 の全ネットワークストレージと、レプリケーション先仮想一元化ファイルシステム 104 に対して転送する。

#### 【0037】

仮想一元化装置 110 は、管理手段 150、要求受付手段 151、レプリケーション手段 152、判定手段 153、マッピング手段 154、ホストテーブル 155、レプリケーションホストテーブル 156、一元管理ディレクトリ 157 で構成する。

#### 【0038】

本実施の形態では、仮想一元化装置 110 はリモート制御手段やストレージ装置持たないが、仮想一元化装置 110 にリモート制御手段やストレージ装置を含んでもよい。

#### 【0039】

管理手段 150 は、仮想一元化装置 110 に常駐し、クライアント 100～1

01からの管理要求に従い、IPアドレス、名前情報、マウントポイント等の基本的な設定情報を設定する。基本的な設定情報は、他に、時刻、アカウント、セキュリティ情報等がある。管理手段150を介したファイルアクセス操作としては、新しく共有ディレクトリを作成したり、ディレクトリのアクセス権限の変更する処理がある。

#### 【0040】

管理手段150は、複数のネットワークストレージを仮想一元化したり、複数の仮想一元化ファイルシステム間でファイルのレプリケーションを行なう場合に、ホストテーブル155、レプリケーションホストテーブル156、一元管理ディレクトリ157の内容を書き換える。また、管理手段150は、レプリケーション手段152に設定される設定情報を書き換える。設定情報の詳細は後述する。さらに、管理手段150は、要求受付手段151、レプリケーション手段152、マッピング手段154を呼び出してファイルアクセス要求を処理する。

#### 【0041】

管理者は、クライアント100～101上の汎用Webブラウザをユーザインタフェースとして管理手段150にアクセスする。あるいは、クライアント上に専用のユーザインタフェースプログラムを搭載してそれを用いてもよい。あるいは、仮想一元化装置110にコンソールを接続してそれを用いてもよい。仮想一元化装置110には、Webブラウザ等を利用してアクセスできるように、Webサーバを搭載する。Webサーバは公知のApache等を使用する。本実施の形態では、管理手段150は、複数ネットワークストレージから仮想一元化ファイルシステムを構成するための設定を行なうと共に、複数の仮想一元化ファイルシステム間におけるファイルレプリケーションのための設定を行なう。

#### 【0042】

要求受付手段151は、クライアント100～101からのマウント要求を処理する。またクライアント100～101からのファイルアクセス要求を受け付け、そのファイルアクセス要求をレプリケーション手段152に転送する。本実施の形態では、クライアント100～101は、前述のNFSプロトコルを用いて、要求受付手段151にアクセスする。仮想一元化装置110とネットワーク



ストレージ 120～122 のファイルシステム間も NFS プロトコルを用いてファイルアクセスを行なうものとする。クライアントが CIFS クライアントである場合には、CIFS プロトコルを用いて、要求受付手段 151 にアクセスする。その場合、仮想一元化装置 110 の要求受付手段 151 とネットワークストレージ 120～122 のストレージ装置間でのファイルアクセスは、NFS プロトコルを使用する。CIFS プロトコルを使用したファイルアクセスは、Samba 等の公知の技術を用いて実現できるため、ここでは詳細には説明しない。

#### 【0043】

レプリケーション手段 152 は、判定手段 153 を含み、要求受付手段 151 により呼び出される。レプリケーション手段 152 は、一元管理ディレクトリ 157 にアクセスし、判定手段 152 により、要求受付手段 151 から転送されたファイルアクセス要求をレプリケーションするかどうかを判定する。レプリケーションすると判定した場合には、ファイルアクセス要求の対象であるレプリケーション元の仮想一元化ファイルシステム 103 に属するいずれかのネットワークストレージと、レプリケーション先仮想一元化ファイルシステム 104 の仮想一元化装置 111 に対して同時にそのファイルアクセス要求を転送する。判定手段 152 の詳細については後述する。また、一元管理ディレクトリ 157 に関しても後述する。

#### 【0044】

ホストテーブル 155 は、仮想一元化ファイルシステム 103 のメンバ情報を管理する。図 2 にホストテーブル 155 の構成を示す。ホストテーブル 155 は、仮想一元化ファイルシステム 103 のメンバとなるネットワークストレージの名前情報とメンバ間で一意なネットワークストレージ識別子との対応付けを行なう。201 の行はメンバとなるネットワークストレージのストレージ装置の名前情報であり、202 の行は、各ストレージ装置のマウントポイント、203 の行は各ネットワークストレージのストレージ装置に対応する識別子である。一つのネットワークストレージに複数のストレージ装置が存在する場合には、各ストレージ装置に対応して識別子を用意する。図 2 では、ネットワークストレージ 120 (NS120) のネットワークストレージ識別子を PS1、ネットワークスト

レージ 1 2 1 ( N S 1 2 1 ) のネットワークストレージ識別子を P S 2、ネットワークストレージ 1 2 2 ( N S 1 2 2 ) のネットワークストレージ識別子を P S 3 に設定する。

#### 【 0 0 4 5 】

本実施の形態では、仮想一元化装置 1 1 0 はリモート制御手段やストレージ装置を持たないが、リモート制御手段やストレージ装置を持つ場合には、ホストテーブル 1 5 5 に仮想一元化装置 1 1 0 の名前情報とストレージ装置のマウントポイント、ストレージ装置に対する識別子を登録すればよい。

#### 【 0 0 4 6 】

レプリケーションホストテーブル 1 5 6 は、レプリケーション先の仮想一元化ファイルシステムのメンバ情報を管理する。図 3 にレプリケーションホストテーブル 1 5 6 の構成を示す。レプリケーションホストテーブル 1 5 6 の構成はホストテーブル 1 5 5 と同様である。但し、レプリケーションホストテーブル 1 5 6 は、仮想一元化ファイルシステムに所属する各ネットワークストレージ装置ではなく、仮想一元化ファイルシステムそのものを管理する。3 0 1 の行は、レプリケーションのメンバとなる仮想一元化ファイルシステムの名前情報を管理する。3 0 2 の行は、各仮想一元化ファイルシステムのマウントポイントを管理する。3 0 3 の行は、各仮想一元化ファイルシステムの識別子である。図 3 では、仮想一元化ファイルシステム 1 0 4 ( V F S 1 0 3 ) の識別子を V F S 1 に設定している。ファイル及びディレクトリのコピーを一つ以上の仮想一元化ファイルシステムに作成する場合には、レプリケーションホストテーブル 1 5 6 に追加すればよい。本発明の実施の形態では、レプリケーションホストテーブル 1 5 6 でレプリケーション先の仮想一元化ファイルシステムを管理するが、ホストテーブル 1 5 5 にレプリケーション先の仮想一元化ファイルシステムの名前情報とマウントポイントを含んでもよい。

#### 【 0 0 4 7 】

一元管理ディレクトリ 1 5 7 は、ネットワークストレージ 1 2 0 ~ 1 2 2 を含む仮想一元化ファイルシステム 1 0 3 全体のファイルディレクトリツリー構造と、全ファイル識別子 ( 属性情報 ) のみを管理する。クライアントには一元管理デ

ィレクトリ 157 のファイルディレクトリツリー構造を見せるが、一元管理ディレクトリ 157 上の各ファイルは、各ファイルの実体であるデータを持たない。全データは、ネットワークストレージ 120～122 に分散して格納する。

#### 【0048】

本実施の形態では、クライアントには仮想的に単一のファイルディレクトリツリー構造を見せるための一元管理ディレクトリ 157 として、ファイルシステムを利用する。一般的なファイルシステムは、ファイルディレクトリツリー構造とファイル識別子と共に、ファイルの実体であるデータも格納してしまう。そこで、仮想一元化装置 110 は、一元管理ディレクトリ 157 のファイルシステムにダミーファイルを生成する。ダミーファイルはファイルの実体であるデータを持たない。

#### 【0049】

クライアントが仮想一元化装置 110 にファイル生成要求 (NFS プロトコルでは CREATE 要求に相当する) を発行すると、レプリケーション手段 152 は一元管理ディレクトリ 157 にファイル構造を作るが、ファイルの実体はネットワークストレージ 120～122 のファイルシステムに格納する。また、クライアント 100 が仮想一元化装置 110 に READ 要求を発行すると、要求受付手段 151 が READ 要求を受け付け、レプリケーション手段 152 に転送する。レプリケーション手段 152 は、一元管理ディレクトリ 157 にアクセスして、ダミーファイルを特定し、ダミーファイルを利用してファイルの実体を格納するネットワークストレージを決定する。次に、そのネットワークストレージからファイルの実体であるデータを読み出し、一元管理ディレクトリ 157 で管理する属性情報とあわせて、クライアントに返送する。なお、一元管理ディレクトリ 157 は、ファイルディレクトリツリー構造とファイル識別子の管理さえできればよいので、必ずしもファイルシステムを使用する必要はない。ファイル識別子が全ファイルに対して一意であることを保証できれば、データベースを利用してよいし、専用のテーブルで管理してもよい。

#### 【0050】

マッピング手段 154 は、ファイルとファイルの格納先のネットワークストレ

ージとを対応つける。本実施の形態では、マッピング手段 1 5 4 は、一元管理ディレクトリ 1 5 7 に作成したダミーファイルの中に、ファイルの実体を格納するネットワークストレージの識別子を格納する。ネットワークストレージの識別子はホストテーブル 1 5 5 に登録したシステムに一意な値であり、各ダミーファイルに書きこむネットワークストレージの識別子は、各ネットワークストレージに均等にファイルが分配されるように決定する。具体的には、CREATE 要求によりファイルを新しく生成する場合に、ラウンドロビンでファイルの実体を格納するネットワークストレージの識別子を決定する。その他にも、各ネットワークストレージの容量を監視して、格納されている容量が均等になるようにネットワークストレージを決定することもできる。あるいは、特定ディレクトリ以下のファイルを全て同じネットワークストレージに割り当てることもできる。

#### 【0 0 5 1】

本実施の形態では、ダミーファイルにファイルの実体を格納するネットワークストレージの識別子を格納する方法をとったが、ダミーファイルではなく、各ディレクトリ毎に保持するファイル名とファイルの実体を格納するネットワークストレージの識別子を対応付けるリストを別途保持してもよい。ファイル識別子から、各ファイルの実体を格納するネットワークストレージが一意に決定できる方法であれば、専用のテーブルで管理してもよいし、他のどのような方法を用いてもよい。

#### 【0 0 5 2】

次に、レプリケーション手段 1 5 2 について図 4 を用いて詳細に説明する。レプリケーション手段 1 5 2 は、判定手段 1 5 3、要求処理手段 4 0 1、応答収集手段 4 0 2、整合手段 4 0 3、同期フラグ 4 1 0、マスタフラグ 4 1 1、ルールテーブル 4 1 2 で構成する。

#### 【0 0 5 3】

判定手段 1 5 3 は、要求受付手段 1 5 1 から呼び出され、要求受付手段 1 5 1 が受け付けたファイルアクセス要求をレプリケーションするかどうかを判定する。本実施の形態では、要求受付手段 1 5 1 が受け付けたファイルアクセス要求が更新系要求であること、同期フラグ 4 1 0 が設定されていること、マスタフラグ

411が設定されていること、ファイルアクセス要求がコピー対象であること、の4つのAND条件が成立した場合にレプリケーションを行なう。同期フラグ410およびマスタフラグ411に関しては後述する。ファイルアクセス要求が参照系要求である場合には、レプリケーション先仮想一元化ファイルシステム104にはそのアクセス要求を転送しない。また、同期フラグの値が“偽”である場合、レプリケーション先仮想一元化ファイルシステム104にはその要求を転送しない。さらに、マスタフラグ411の値が“偽”である場合にも、レプリケーション先仮想一元化ファイルシステム104にはアクセス要求を転送しない。その上、ファイルアクセス要求がコピー対象でなければ、レプリケーション先仮想一元化ファイルシステム104には、ファイルアクセス要求を転送しない。

#### 【0054】

要求処理手段401は、判定手段153の判定の結果、レプリケーションを行うと判定した場合に、要求受付手段151が受け付けたファイルアクセス要求をレプリケーション元仮想一元化ファイルシステム103に属するネットワークストレージ120～122と、レプリケーション先仮想一元化ファイルシステム104の仮想一元化装置111に同時に転送する。判定手段153の判定の結果、レプリケーションしないと判定した場合には、要求受付手段151が受け付けたファイルアクセス要求をレプリケーション元仮想一元化ファイルシステム103に属するネットワークストレージ120～122にのみ転送する。

#### 【0055】

応答収集手段402は、要求処理手段401がレプリケーション元仮想一元化ファイルシステム103に属するネットワークストレージ120～122と、レプリケーション先仮想一元化ファイルシステム104からの応答をまとめて一つにしてクライアントに返送する。

#### 【0056】

整合手段403は、図1に示す管理手段150により呼び出され、レプリケーション元仮想一元化ファイルシステム103に存在するコピー対象のファイル及びディレクトリとレプリケーション先仮想一元化ファイルシステム104に存在するファイルおよびディレクトリの整合性をとる。具体的には、コピー対象のフ

ファイルおよびディレクトリが一致するかどうかを検査し、一致しない場合には、レプリケーション元仮想一元化ファイルシステム 1 0 3 の内容を最新の情報として、レプリケーション先仮想一元化ファイルシステム 1 0 4 に反映する。

#### 【0 0 5 7】

同期フラグ 4 1 0 は、整合手段 4 0 3 により設定され、レプリケーション元仮想一元化ファイルシステム 1 0 3 におけるコピー対象のファイル及びディレクトリと、レプリケーション先仮想一元化ファイルシステム 1 0 4 におけるコピーしたファイル及びディレクトリが一致するかどうかを示す。同期フラグ 4 1 0 の初期化時は“偽”と設定する。一致している場合には“真”とし、不一致の場合には“偽”とする。

#### 【0 0 5 8】

マスタフラグ 4 1 1 は、管理手段 1 5 0 により設定され、レプリケーション元仮想一元化ファイルシステム 1 0 3 のファイル及びディレクトリがマスタであることを示す。マスタフラグ 4 1 1 は、初期化時に“偽”と設定する。レプリケーション元仮想一元化ファイルシステム 1 0 3 のレプリケーション手段 1 5 2 のマスタフラグ 4 1 1 には、“真”と設定し、レプリケーションホストテーブル 1 5 6 に登録したレプリケーション先仮想一元化ファイルシステム 1 0 4 のレプリケーション手段 1 6 2 のマスタフラグ 4 1 1 には“偽”と設定する。

ルールテーブル 4 1 2 は、管理手段 1 5 0 により設定され、コピー対象の特定ディレクトリ全体あるいは特定ユーザ識別子あるいは特定グループ識別子を保持する。あるいは、コピー対象ファイルのファイルの識別子を保持する。例えば、管理者が特定ユーザのファイルをコピーすると設定する場合には、ルールテーブル 4 1 2 にそのユーザの識別子を格納する。レプリケーション元仮想一元化ファイルシステム 1 0 3 に属する全ファイルのコピーを作成する場合には、ルールテーブル 4 1 2 のかわりに、コピーフラグを用意すればよい。あるいは、ルールテーブル 4 1 2 に全ファイルディレクトリのコピーを作成するルールを記載してもよい。あるいは、ルールテーブル 4 1 2 とコピーフラグを併用してもよい。

#### 【0 0 5 9】

ファイルの識別子によりファイルのコピーを保持するレプリケーション先の仮

想一元化ファイルシステムを変更する場合には、ルールテーブル 4 1 2 のルール毎に、レプリケーションホストテーブル 1 5 6 に設定したレプリケーション先の仮想一元化ファイルシステムの識別子を対応させればよい。これにより、レプリケーション先の仮想一元化ファイルシステムが複数存在する場合に、仮想一元化ファイルシステム毎に異なる条件でファイルのコピーを作成することができる。さらに、ファイル識別子毎にレプリケーション先仮想一元化ファイルシステムを変更する場合には、レプリケーションホストテーブル 1 5 6 の各メンバに対応する同期フラグを持つことにより、より細かな条件でファイルシステムの部分コピーを作成することができる。

#### 【 0 0 6 0 】

本実施の形態では、ルールテーブル 4 1 2 にコピー条件を保持する方法をとったが、ルールテーブル 4 1 2 ではなく、一元管理ディレクトリ 1 5 7 の各ダミーファイルにコピーするかどうかを示すコピー識別子を保持してもよい。あるいは、各ディレクトリ毎にディレクトリ内に保持するファイルとコピー識別子を対応付けるリストを別途保持してもよい。ファイル識別子から、そのファイルがコピー対象であるかどうかを一意に決定できる方法であれば、専用のテーブルで管理してもよいし、他のどのような方法を用いてもよい。

#### 【 0 0 6 1 】

レプリケーション先仮想一元化ファイルシステム 1 0 4 の構成は、レプリケーション元仮想一元化ファイルシステム 1 0 3 と同じであるため、詳細説明は省略する。ただし、本実施の形態では、レプリケーション先仮想一元化ファイルシステム 1 0 4 のレプリケーションホストテーブル 1 6 6 にはレプリケーション先の仮想一元化ファイルシステムは登録されていない。また、レプリケーション手段 1 6 2 が使用する同期フラグおよびマスタフラグの値も“偽”に設定される。これらの設定は、レプリケーション元仮想一元化ファイルシステム 1 0 3 の管理手段 1 5 0 を利用して設定してもよいし、管理者が直接レプリケーション先仮想一元化ファイルシステム 1 0 4 の管理手段 1 6 0 にアクセスして設定してもよい。

#### 【 0 0 6 2 】

次に、レプリケーション手段 1 5 2 が行なうレプリケーション処理について図

5を用いて詳細に説明する。要求受付手段151がクライアントからファイルアクセス要求を受け付けると、要求受付手段151はレプリケーション手段152を呼び出し、レプリケーション処理を行なう。

#### 【0063】

レプリケーション処理は、判定処理510、要求転送処理511、応答収集処理512からなる。

#### 【0064】

判定処理510は、プロセス502において、要求受付手段151が受け付けたファイルアクセス要求をレプリケーションするかどうかを判定する。レプリケーションをする場合には要求転送処理411のプロセス504を、レプリケーションしない場合にはプロセス503を行なう。判定処理510の詳細については後述する。

#### 【0065】

要求転送処理511は、プロセス502の判定結果に従って、プロセス503あるいはプロセス504を行なう。プロセス503は、判定処理510の判定結果が“YES”である場合に行なわれ、ファイルアクセス要求をレプリケーション元仮想一元化ファイルシステム103のホストテーブル155に登録したファイルの実体を格納するネットワークストレージに転送すると共に、レプリケーションホストテーブル156に登録した仮想一元化ファイルシステム104に対してもそのファイルアクセス要求を転送する。本実施の形態では、ディレクトリに対するファイルアクセス要求の場合、レプリケーション元仮想一元化ファイルシステム103に属する全ネットワークストレージに対してディレクトリに対するファイルアクセス要求を転送すると共に、レプリケーションホストテーブル156に登録した仮想一元化ファイルシステム104に転送する。

#### 【0066】

プロセス504は、判定処理510の判定結果が“NO”である場合に行なわれ、ファイルアクセス要求をレプリケーション元仮想一元化ファイルシステム103のホストテーブル155に登録したファイルの実体を格納するネットワークストレージのみに転送する。ディレクトリに対するファイルアクセス要求の場合



、レプリケーション元仮想一元化ファイルシステム 1 0 3 に属する全ネットワークストレージに対してアクセス要求を転送する。

#### 【0 0 6 7】

応答収集処理 5 1 2 は、要求転送処理 5 1 1 の応答を収集し、一つにしてクライアントに返送する。プロセス 5 0 5 は、判定処理 5 1 0 の判定結果が“YES”の場合に行なわれ、ファイルに対するアクセス要求の場合には、レプリケーション元仮想一元化ファイルシステム 1 0 3 のファイル格納先ネットワークストレージからの応答と、レプリケーション先仮想一元化ファイルシステム 1 0 4 からの応答をあわせて一つの応答にしてクライアント 1 0 0 に返送する。ディレクトリに対するファイルアクセス要求の場合、レプリケーション元仮想一元化ファイルシステム 1 0 3 の全てのネットワークストレージからの応答と、レプリケーション先仮想一元化ファイルシステム 1 0 4 からの応答をあわせて一つにしてクライアント 1 0 0 に返送する。

#### 【0 0 6 8】

プロセス 5 0 6 は、判定処理 5 1 0 の判定結果が“NO”の場合に行なわれ、ファイルに対するファイルアクセス要求の場合、レプリケーション元仮想一元化ファイルシステム 1 0 3 のファイル格納先ネットワークストレージからの応答をクライアント 1 0 0 に返送する。ディレクトリに対するファイルアクセス要求の場合には、レプリケーション元仮想一元化ファイルシステム 1 0 3 の全てのネットワークストレージからの応答をまとめて一つにしてクライアント 1 0 0 に返送する。

#### 【0 0 6 9】

本実施の形態では、応答収集処理 5 1 2 が転送先のネットワークストレージや仮想一元化ファイルシステムからの応答全てを待ち合わせてからクライアントに応答を返送する例を示したが、クライアントに対するレスポンス時間を短縮するため、複数の転送先からの最も早い応答を受け取った後、他の応答を待たずにクライアント 1 0 0 にその応答を返送してもよい。

#### 【0 0 7 0】

次に、判定処理 5 1 0 の詳細フローについて図 1、図 4、および図 6 を用いて

説明する。要求受付手段 1 5 1 がファイルアクセス要求を受け付けると、判定手段 1 5 3 を呼び出す。判定手段 1 5 3 は、判定処理 5 1 0 を行なう。判定処理 5 1 0 は、プロセス 6 0 2 においてファイルアクセス要求が更新系要求であるかどうかを判定する。更新系要求である場合には、判定結果を“Y E S”としてプロセス 6 0 3 を行なう。

#### 【0 0 7 1】

プロセス 6 0 3 では、マスタフラグ 4 1 1 を調査し、コピー対象のファイル及びディレクトリがオリジナルであるかどうかを調査する。マスタフラグ 4 1 1 の値が“真”である場合には、判定結果は“Y E S”となり、プロセス 6 0 4 を行なう。

#### 【0 0 7 2】

プロセス 6 0 4 では、同期フラグ 4 1 0 により、コピー対象のファイル及びディレクトリとレプリケーション先仮想一元化ファイルシステムのコピー先ファイルおよびディレクトリが一致するかどうかを調査する。同期フラグ 4 1 0 の値が“真”である場合には、判定結果は“Y E S”となり、プロセス 6 0 5 を行なう。

#### 【0 0 7 3】

プロセス 6 0 5 は、ファイルアクセス要求のアクセス対象のファイルオブジェクトがコピー対象であるかどうかを判定する。プロセス 6 0 5 は、プロセス 6 0 9、プロセス 6 1 0 からなる。プロセス 6 0 9 では、ファイルアクセス要求のアクセス対象のファイルオブジェクトがディレクトリである場合には、コピー対象として判定結果を“Y E S”にして、プロセス 6 0 6 を行う。ファイルアクセス要求の対象ファイルオブジェクトが、ファイルである場合には、プロセス 6 1 0 を行う。プロセス 6 1 0 は、ファイルアクセス対象ファイルオブジェクトが図 4 に示すルールテーブル 4 1 2 のコピー条件を満たすかどうかを判定する。判定では、一元管理ディレクトリ 1 5 7 にアクセスし、ファイルの属性情報を取得し、取得した属性情報とルールテーブル 4 1 2 のコピー条件を比較する。コピー条件が成立する場合は、判定結果を“Y E S”とし、プロセス 6 0 6 を行う。条件不成立の場合は、コピー対象外ファイルとして判定結果を“N O”としてプロセ

ス 607 を行う。

#### 【0074】

本実施の形態では、ファイルアクセス要求の対象ファイルオブジェクトがディレクトリである場合には、必ずコピー対象とする。そのため、レプリケーション元仮想一元化ファイルシステムとレプリケーション先仮想一元化ファイルシステム間で同じディレクトリ構造を持つ。このようにすることで、障害復旧時の整合処理を簡単にすることができる。ただし、ディレクトリに対するファイルアクセス要求の場合にも、ルールテーブル 412 を利用してコピー条件が成立するかどうかを判定することにより、ファイルだけでなく、ディレクトリに関してもコピー対象かどうかを個別に選択する方法も容易に実現できる。また、コピーフラグを用いて、仮想一元化ファイルシステム 103 のファイルシステム全体をコピーする場合、プロセス 605 において、ファイルやディレクトリの区別なく、コピーフラグの値が“真”かどうかを調査することにより、コピー判定を行えばよい。

#### 【0075】

プロセス 606 では、レプリケーション判定を“YES”として判定処理 510 を完了する。プロセス 602、プロセス 603、プロセス 604、プロセス 605 のいずれかの判定結果が“NO”である場合には、プロセス 607 を行ない、レプリケーション判定を“NO”として判定処理 510 を完了する。

#### 【0076】

次に、レプリケーション判定が“YES”である場合の、要求転送処理 511 及び応答収集処理 512 における、ファイルアクセス要求の処理について詳細に説明する。本実施の形態では、レプリケーション手段 152 が受け付けたファイルアクセス要求に関して、判定処理 510 において、判定結果が“YES”となる場合に、そのアクセス要求を、レプリケーション元仮想一元化ファイルシステム 103 だけでなく、レプリケーション先仮想一元化ファイルシステム 104 にも転送する。ファイルアクセス要求の処理には NFS プロトコルを利用する。ここで、NFS バージョン 2 における更新系要求は、CREATE 要求、WRITE 要求、REMOVE 要求、RENAME 要求、MKDIR 要求、RMDIR 要

求、SETATTR要求、LINK要求、SYMLINK要求である。

#### 【0077】

最初に、ファイルハンドルの構成について図7を用いて説明する。ファイルハンドル700は親ディレクトリのiノード番号フィールド701と、ファイルオブジェクトのiノード番号フィールド702を含む。

#### 【0078】

次に、本実施の形態において、レプリケーションの対象となる更新系要求処理の詳細について説明する。

#### 【0079】

図8にCREATE要求処理の流れを示す。CREATE要求は、ファイルを生成するディレクトリのファイルハンドルHと、ファイルの名前Fを引数とする。図1のレプリケーション手段152は、要求受付手段151からCREATE要求を受け取るとプロセス801を呼び出す。

#### 【0080】

プロセス802では、ファイルハンドルHのフィールド702から、ファイルを生成するディレクトリのiノード番号PIを取りだし、このiノード番号で識別されるディレクトリにファイルの名前がFとなるダミーファイルを生成する。

プロセス803では、マッピング手段154およびホストテーブル155を利用して、CREATE要求のファイルを格納するネットワークストレージ識別子PSを決定する。

#### 【0081】

プロセス804では、識別子PSをダミーファイルFに書き込む。書きこみは、例えばWRITE要求を利用する。

#### 【0082】

プロセス805では、レプリケーションホストテーブル156を利用して、CREATE要求を転送するレプリケーション先仮想一元化ファイルシステムVFSを決定する。

#### 【0083】

プロセス 806 では、クライアントから送られてきた CREATE 要求をネットワークストレージ PS と仮想一元化ファイルシステム VFS に対して発行する。

#### 【0084】

プロセス 807 では、ネットワークストレージ PS と、仮想一元化ファイルシステム VFS の両方から、CREATE 要求に対する応答を収集する。

#### 【0085】

プロセス 808 では、ダミーファイル F を生成したディレクトリの i ノード番号 PI とダミーファイルの i ノード番号 I からファイルハンドル 700 を構成して、クライアント 100 に返す。

#### 【0086】

図 9 は WR I T E 要求処理の流れを示す。WR I T E 要求は、読み出すファイル F のファイルハンドル H と読み出すファイル F の親ディレクトリのファイルハンドル H p を引数とする。図 1 に示すようにレプリケーション手段 152 が要求受付手段 151 から WR I T E 要求を受け取ると、図 9 のプロセス 901 を呼び出し、WR I T E 要求処理を行なう。

#### 【0087】

プロセス 902 では、ファイルハンドル H からダミーファイル F を読み出し、ダミーファイル F からファイル格納先のネットワークストレージ PS を求める。

#### 【0088】

プロセス 903 では、ネットワークストレージ PS からファイル F の実体に対するファイルハンドル H p s を取得する。ファイルハンドル H p s は、ネットワークストレージ PS に対して L O O K U P 要求を発行することにより取得できる。一度取得したファイルハンドル H p s の値とファイルハンドル H との対応関係を仮想一元化装置 110 に記録しておくことで次にファイルハンドル H を用いて WR I T E 要求を発行する場合には、L O O K U P 要求を発行しなくてもファイル格納先ネットワークストレージ PS におけるファイルハンドル H p s を取得できる。

#### 【0089】

プロセス 9 0 4 では、レプリケーションホストテーブル 1 5 6 に登録した仮想一元化ファイルシステム V F S から、ファイル F の実体に対するファイルハンドル H v f s を取得する。ファイルハンドル H v f s は、仮想一元化ファイルシステム V F S に対して L O O K U P 要求を発行することにより取得できる。一度取得したファイルハンドル H v f s の値とファイルハンドル H との対応関係を仮想一元化装置 1 1 0 に記録しておくことで、次にファイルハンドル H を用いて仮想一元化ファイルシステム V F S に対して W R I T E 要求を発行する場合に、L O O K U P 要求を発行しなくてもレプリケーション先仮想一元化ファイルシステム V F S におけるファイルハンドル H v f s を取得できる。

#### 【 0 0 9 0 】

プロセス 9 0 5 では、取得したファイルハンドル H p s を引数として W R I T E 要求をネットワークストレージ P S に発行する。さらに、取得したレプリケーション先のファイルハンドル H v f s を引数として W R I T E 要求をレプリケーション先仮想一元化ファイルシステム V F S に発行する。

#### 【 0 0 9 1 】

プロセス 9 0 6 では、ネットワークストレージ P S および仮想一元化ファイルシステム V F S それぞれが W R I T E 要求の返り値を返すと、その結果をまとめて一つにしてクライアント 1 0 0 に返す。

#### 【 0 0 9 2 】

R E M O V E 要求処理と R E N A M E 要求処理も W R I T E 要求処理と同じであるため、詳細説明は省略する。

#### 【 0 0 9 3 】

図 1 0 に M K D I R 要求処理の流れを示す。M K D I R 要求は、生成するディレクトリの名前 D と、生成する親ディレクトリのファイルハンドル H を引数とする。図 1 に示すように、レプリケーション手段 1 5 2 が要求受付手段 1 5 1 から M K D I R 要求を受け付けると、図 1 0 のプロセス 1 0 0 1 を行なう。

#### 【 0 0 9 4 】

プロセス 1 0 0 2 では、ファイルハンドル H のフィールド 7 0 2 に保持された i ノード番号によって識別される一元管理ディレクトリ 1 5 7 のディレクトリ内

に、名前Dのディレクトリを生成する。生成したディレクトリのiノード番号はIとする。

#### 【0095】

プロセス1003では、ホストテーブル155に登録されている全ネットワークストレージと、レプリケーションホストテーブル156に登録されている全仮想一元化ファイルシステムに対して、MKDIR要求を発行して名前Dのディレクトリを生成する。

#### 【0096】

プロセス1004では、プロセス1003で転送したMKDIR要求からの応答全てを受け取ると、ディレクトリDの親ディレクトリのiノード番号PIと一元管理ディレクトリ156に生成したディレクトリDのiノード番号Iからファイルハンドルを生成し、クライアント100に返送する。

RMDIR要求処理もMKDIR要求処理と同じであり、MKDIR要求をRMDIR要求に変更すれば容易に実現できる。SYMLINK要求処理もMKDIR要求処理と同じであり、MKDIR要求をSYMLINK要求に変更すればよい。SETATTR要求に関しては、アクセス対象のファイルオブジェクトがディレクトリの場合、MKDIR要求処理と同様に、各ネットワークストレージとレプリケーション先の仮想一元化ファイルシステムに対してSETATTR要求を発行してディレクトリ属性の設定を行なう。なお、アクセス対象のファイルオブジェクトがファイルであるSETATTR要求処理は、WRITE要求処理と同様に実現できる。

#### 【0097】

本実施の形態では、クライアントからのファイルアクセス要求が参照系要求である場合、図6に示す判定処理の結果は“NO”となり、レプリケーションを行なわない。NFSバージョン2における参照系要求としては、LOOKUP要求、READ要求、READLINK要求、REaddir要求、GETATTR要求、STATFS要求がある。

#### 【0098】

図11にLOOKUP要求処理の流れを示す。LOOKUP要求は、ファイル

オブジェクトの親ディレクトリのファイルハンドルHと、ファイルオブジェクトの名前Fを引数とする。図1に示すようにレプリケーション手段152が、要求処理手段151からLOOKUP要求を受け取ると、図11のプロセス1101を呼び出す。

#### 【0099】

プロセス1102において、ファイルハンドルHからフィールド201に保持されたiノード番号PIを取りだし、このiノード番号で識別される一元管理ディレクトリ157に存在する名前Fのファイルオブジェクトのiノード番号Iを取得する。次にプロセス1103において、親ディレクトリのiノード番号PIとファイルオブジェクトFのiノード番号Iをあわせてファイルハンドルを構成し、それをクライアント100に返す。

READDIR要求の処理はLOOKUP要求と同様で、対象ディレクトリの情報を一元管理ディレクトリ157から読み出し、クライアント100に返す。

#### 【0100】

図12はREAD要求処理の流れを示す。READ要求は、読み出すファイルFのファイルハンドルHと読み出すファイルFの親ディレクトリのファイルハンドルHpを引数とする。図1に示すように、レプリケーション手段152が要求受付手段151からREAD要求を受け取ると、図12のプロセス1201を呼び出す。プロセス1202では、ファイルハンドルHからダミーファイルFを読み出し、ダミーファイルFからファイル実体格納先のネットワークストレージPSを求める。プロセス1203では、ネットワークストレージPSからファイルFの実体に対するファイルハンドルHpsを取得する。ファイルハンドルHpsは、ネットワークストレージPSに対してLOOKUP要求を発行することにより取得できる。一度取得したファイルハンドルHpsの値とファイルハンドルHとの対応関係を一元管理ディレクトリ157内部に記録しておくことで次にファイルハンドルHを用いてREAD要求を発行する場合には、LOOKUP要求を発行しなくてもファイル格納先ネットワークストレージPSにおけるファイルハンドルHpsを取得できる。

#### 【0101】



プロセス 1 2 0 4 では、取得したファイルハンドル `H p s` を引数として `R E A D` 要求をネットワークストレージ `P S` に発行する。プロセス 1 2 0 5 では、ネットワークストレージ `P S` が `R E A D` 要求の返り値である読み出し結果を返すと、その結果をクライアント 1 0 0 に返す。

`G E T A T T R` 要求は、ホストテーブル 1 5 5 に登録した各ネットワークストレージに対して要求を発行してディレクトリ属性の読み出しを行なう。アクセス対象のファイルオブジェクトがファイルの場合、図 1 2 に示す `R E A D` 要求処理と同様ファイル格納先のネットワークストレージに発行すればよい。

### 【 0 1 0 2 】

次に、図 4 に示す整合手段 4 0 3 を用いた整合処理の流れについて詳細に説明する。図 1 に示すようにレプリケーション元仮想一元化ファイルシステム 1 0 3 とレプリケーション先仮想一元化ファイルシステム 1 0 4 でファイルのレプリケーションを行なっている場合に、どちらかのファイルシステムに障害が発生したり、ネットワーク等に一時的な障害が発生すると、同期フラグ 4 1 0 の値を“偽”として、レプリケーション処理を停止する。クライアントがレプリケーション元仮想一元化ファイルシステム 1 0 3 にアクセス可能な場合は、仮想一元化ファイルシステム 1 0 3 のみで運用を継続する。仮想一元化ファイルシステム 1 0 3 に障害が発生し、クライアントがアクセスできない場合には、管理者がコピー対象のファイルのみを持つレプリケーション先仮想一元化ファイルシステム 1 0 4 をクライアントに公開するように設定を変更する。例えば、レプリケーション先仮想一元化ファイルシステム 1 0 4 の `e x p o r t s` ファイルを変更する。

### 【 0 1 0 3 】

レプリケーション元仮想一元化ファイルシステム 1 0 3 のみで運用を継続し、ネットワーク等の一時的な障害が復旧し、レプリケーション先仮想一元化ファイルシステム 1 0 4 と再度レプリケーションを行なう場合、管理者は管理手段 1 5 0 を用いて整合手段 4 0 3 を呼び出し、複数の仮想一元化ファイルシステムの間で、コピー対象ファイルおよびディレクトリの整合をとる。

### 【 0 1 0 4 】

整合手段 4 0 3 は、最初仮想一元化ファイルシステム 1 0 3 のみで運用してい

る場合に、新しい仮想一元化ファイルシステム 104 をレプリケーション先として追加して、仮想一元化ファイルシステム 104 にコピー対象のファイルおよびディレクトリ全てをコピーし、あたらしくレプリケーション運用する場合にも用いられる。

### 【0105】

図 13 に整合処理の流れを説明する。管理者が整合要求を管理手段 150 に発行すると、管理手段 150 は整合手段 403 を呼び出し、整合処理を開始する。プロセス 1302 では、一元管理ディレクトリ 157 のファイルディレクトリツリーに未探索のファイルオブジェクト F が存在するかどうかをチェックする。未探索のファイルオブジェクト F が存在する場合は、プロセス 1303 において F がディレクトリであるかどうかを判定する。F がファイルである場合には、プロセス 1304 において、F がルールテーブル 412 に設定したコピー条件を満たすかどうかを調査する。コピー条件が成立しない場合には、コピーの必要はないため、プロセス 1302 に戻る。条件が成立する場合には、プロセス 1305 を行なう。プロセス 1303 において、F がディレクトリである場合にはプロセス 1305 を行なう。プロセス 1305 では、レプリケーションホストテーブル 156 に登録した仮想一元化ファイルシステム 104 (VFS 104) にアクセスし、コピーしたファイルオブジェクト F (コピー F) を読み出す。プロセス 1306 では、コピー F があるかどうかを調査する。もし、VFS 104 にコピー F が存在しなければ、プロセス 1307 において VFS 104 に新しくコピー F を作成し、プロセス 1302 に戻る。ファイルの場合、コピー F の作成は、CREATE 要求と WRITE 要求を VFS 104 に発行することにより実現できる。プロセス 1306 において、コピー F が存在する場合にはプロセス 1308 を行なう。プロセス 1308 では、レプリケーション元仮想一元化ファイルシステムである VFS 103 に存在するファイルオブジェクト F とレプリケーション先仮想一元化ファイルシステム 104 (VFS 104) に存在するコピー F の内容が一致するかどうかを調査する。一致する場合には、プロセス 13302 に戻る。一致しない場合には、プロセス 1309 において、VFS 104 に存在するコピー F に VFS 103 のファイルオブジェクト F を上書きして、プロセス 1302

に戻る。上書きは、V F S 1 0 4 に W R I T E 要求を発行することにより実現できる。プロセス 1 0 3 2 は、未探索のファイルオブジェクト F がなくなれば、プロセス 1 3 1 0 を行なう。プロセス 1 3 1 0 では、V F S 1 0 3 と V F S 1 0 4 において、コピー対象のファイルおよびディレクトリが一致するとして、レプリケーション手段 1 5 2 の同期フラグを“真”に設定して処理を終了する。同期フラグが“真”になったことで、クライアントからのファイルアクセス要求に対して、レプリケーション処理を開始することができる。レプリケーション元仮想一元化ファイルシステム 1 0 3 の全ファイルシステムのコピーを作成する場合、プロセス 1 3 0 3、プロセス 1 3 0 4 を行なわずに、コピーフラグが“真”であるかによりレプリケーション元とレプリケーション先の仮想一元化ファイルシステムの全ファイルおよびディレクトリに対して整合性をとればよい。

#### 【 0 1 0 6 】

本実施の形態では、一元管理ディレクトリ 1 5 7 のファイルディレクトリツリーを探索しレプリケーション元とレプリケーション先の仮想一元化ファイルシステムにおいてファイルオブジェクト F が一致するかどうかを順番に調査する。ファイルオブジェクト毎に順番に整合をとっていくのではなく、最初にレプリケーション元とレプリケーション先の仮想一元化ファイルシステムにおけるファイルディレクトリツリーのリストファイルを取得し、リストファイルを比較することで、レプリケーション元とレプリケーション先の仮想一元化ファイルシステムにおけるコピー対象のファイルおよびディレクトリの整合性を取る方法でもよい。リストを取得するための最も一般的な方法としては、l s コマンドを利用すればよい。

#### 【 0 1 0 7 】

あるいは、レプリケーション元仮想一元化ファイルシステム 1 0 3 の一元管理ディレクトリ 1 5 7 がファイルあるいはディレクトリをレプリケーション先にコピーしたかどうかのコピー完了情報を保持し、最初にレプリケーション先仮想一元化ファイルシステム 1 0 4 にディレクトリツリーのコピーを作成し、その後、コピー完了情報に従って、レプリケーション元からレプリケーション先の仮想一元化ファイルシステムにファイルをコピーする方法でもよい。

**【0 1 0 8】**

次に、本実施の形態におけるレプリケーション処理全体の流れについて説明する。管理者が仮想一元化ファイルシステム 1 0 3 と仮想一元化ファイルシステム 1 0 4 を用意し、仮想一元化ファイルシステム 1 0 3 の部分コピーを仮想一元化ファイルシステム 1 0 4 に生成することを想定する。

**【0 1 0 9】**

図 1 4 に管理者による複数仮想一元化ファイルシステム間のレプリケーション設定手順を示す。レプリケーション設定手順は、仮想一元化ファイルシステムの設定（ステップ 1 4 0 2）、レプリケーション先仮想一元化ファイルシステムの設定（ステップ 1 4 0 3）、レプリケーション対象ファイル設定（ステップ 1 4 0 4）、レプリケーション開始コマンド入力（ステップ 1 4 0 5）の順に行なう。

**【0 1 1 0】**

ステップ 1 4 0 2 では、管理者は、Web ブラウザ経由あるいは、専用のユーザインタフェースプログラムを利用して、レプリケーション元仮想一元化ファイルシステム 1 0 3 とレプリケーション先仮想一元化ファイルシステム 1 0 4 の管理手段 1 5 0、1 6 0 にアクセスして、仮想一元化ファイルシステムの設定を行なう。設定情報としては、仮想一元化装置の IP アドレス、名前情報、仮想一元化装置メンバの IP アドレス、名前情報等がある。管理手段 1 5 0 および 1 6 0 は、ホストテーブル 1 5 5 および 1 6 5 に各仮想一元化ファイルシステムのメンバを登録する。ホストテーブル 1 5 5 には、ネットワークストレージ 1 2 0、1 2 1、1 2 2 を登録し、それぞれに対して識別子 PS 1、PS 2、PS 3 を設定する。またレプリケーション先仮想一元化ファイルシステム 1 0 4 のホストテーブル 1 6 5 には、ネットワークストレージ 1 2 3、1 2 4、1 2 5 を登録し、それぞれに対して、PS 4、PS 5、PS 6 を設定する。管理手段 1 5 0 および 1 6 0 は、メンバ登録処理において、各ネットワークストレージのファイルシステムに一元管理ディレクトリ 1 5 7 および 1 6 7 のディレクトリツリーをコピーする。

**【0 1 1 1】**

ステップ 1 4 0 3 では、管理者は、We b ブラウザ経由あるいは、専用のユーザインタフェースプログラムを利用して、レプリケーション元仮想一元化ファイルシステム 1 0 3 の管理手段 1 5 0 にアクセスして、レプリケーション先仮想一元化ファイルシステム 1 0 4 を設定する。管理手段 1 5 0 は、図 1 に示すレプリケーションホストテーブル 1 5 6 にレプリケーション先仮想一元化ファイルシステム 1 0 4 を登録する。具体的には、レプリケーション先仮想一元化ファイルシステム 1 0 4 の識別子を V F S 1 と設定する。さらに、レプリケーション元仮想一元化ファイルシステム 1 0 3 の管理手段 1 5 0 は、マスタフラグ 4 1 1 を “真” にする。

#### 【0 1 1 2】

ステップ 1 4 0 4 では、ステップ 1 4 0 3 と同様、管理者はレプリケーション元仮想一元化ファイルシステム 1 0 3 にアクセスして、コピー対象のファイルを設定する。例えば、管理者が特定グループに所属するファイルのみをコピー対象とする場合には、ファイル識別子に含まれるグループ識別子を We b ブラウザ経由で、管理手段 1 5 0 に送信する。管理手段 1 5 0 は、その設定内容を、図 4 に示すルールテーブル 4 1 2 に保持する。例えば、本実施の形態では、ルールテーブル 4 1 2 にグループ識別子 = “1” を保持する。

#### 【0 1 1 3】

ステップ 1 4 0 5 では、ステップ 1 4 0 4 と同様、管理者はレプリケーション元仮想一元化ファイルシステム 1 0 3 にアクセスして、レプリケーション開始コマンドを入力する。入力後、レプリケーション開始コマンドが、レプリケーション元仮想一元化ファイルシステム 1 0 3 の管理手段 1 5 0 に書きこまれる。書き込みは、HTML プロトコルの P U T コマンドを利用して実現できる。そのほかにも、We b 対応のインタフェースを利用すれば公知の技術で実現できる。管理手段 1 5 0 は、管理者からの書き込みを契機として、レプリケーション元仮想一元化ファイルシステム 1 0 3 においてレプリケーション処理を開始する。このとき、同期フラグ 4 1 0 が “偽” である場合には、整合手段 4 0 3 を起動して、整合処理を行ない、仮想一元化ファイルシステム 1 0 3 と仮想一元化ファイルシステム 1 0 4 におけるコピー対象ファイルの整合をとる。整合を取った後、同期フ

ラグ 4 1 0 を “真” にする。同期フラグ 4 1 0 が “真” である場合には、仮想一元化ファイルシステム 1 0 3 と仮想一元化ファイルシステム 1 0 4 の間でファイルのレプリケーション処理を開始することができる。

#### 【0 1 1 4】

次に、ファイルレプリケーション処理の例を示す。本実施の形態では、クライアントとレプリケーション元仮想一元化ファイルシステム 1 0 3 間のファイルアクセスでは NFS プロトコルを利用するものとする。NFS プロトコルでは、最初にクライアントがレプリケーション元仮想一元化ファイルシステム 1 0 3 に対してマウント要求を発行して、共有ファイルのマウントポイントをクライアントのファイルシステムの一部にマウントする。レプリケーション元仮想一元化ファイルシステム 1 0 3 の要求受付手段 1 5 1 は、クライアントからのマウント要求に対して、公知の `mountd` を利用してマウント要求に応える。`mountd` は、レプリケーション元仮想一元化ファイルシステム 1 0 3 のマウントポイントのファイルハンドルをクライアントに返送する。以降、クライアントは、ファイルハンドルを利用して、仮想一元化ファイルシステム 1 0 3 にアクセスする。

#### 【0 1 1 5】

例えば、グループ 1 に所属するクライアントが新しくファイルを作成した場合を考える。クライアントが新規ファイル作成要求（NFS プロトコルでは `CREATE` 要求）をレプリケーション元仮想一元化ファイルシステム 1 0 3 に発行すると、要求受付手段 1 5 1 がその要求を受け付ける。`CREATE` 要求は、ファイル識別子にグループ識別子 = “1”、ユーザ識別子 = “5” を含む。要求受付手段 1 5 1 は、次にレプリケーション手段 1 5 2 を呼び出す。レプリケーション手段 1 5 2 は、図 5 に示すレプリケーション処理を開始する。まずプロセス 5 1 0 において、判定処理を行なう。図 6 に示すように、判定処理 5 1 0 では、プロセス 6 0 2 において、`CREATE` 要求は更新系要求であるため、判定結果は “YES” となる。プロセス 6 0 3 では、レプリケーション元仮想一元化ファイルシステム 1 0 3 のマスタフラグ 4 1 1 は “真” であるため、判定結果は “YES” となる。さらにプロセス 6 0 4 において、レプリケーション元仮想一元化ファイルシステム 1 0 3 の同期フラグ 4 1 0 は “真” であるため判定結果は “YES”

”となる。次にプロセス609において、CREATE要求はファイルを対象としているため、プロセス610を行ない、図4に示すルールテーブル412の条件が成立するかどうかを判定する。先に記載したように、ルールテーブル412には、グループ識別子=“1”が保持されている。プロセス610では、CREATE要求のグループ識別子とルールテーブル412に保持したグループ識別子と比較する。比較結果、条件が成立するため、プロセス606においてレプリケーション判定は“YES”となり、判定処理510を終了する。

#### 【0116】

次に図5に示す要求転送処理511を行なう。レプリケーション判定結果は“YES”であるため、プロセス503を行なう。プロセス503では、具体的に、図8に示すCREATE要求処理を実行する。CREATE要求処理では、まずプロセス802において、レプリケーション元仮想一元化ファイルシステム157にダミーファイルを生成し、そのファイルのiノード番号を取得する。プロセス803では、ホストテーブル155とマッピング手段154を利用して、ファイル格納先のネットワークストレージPSを決定する。ここで、ファイル格納先のネットワークストレージは、PS1とする。次にプロセス804において、ダミーファイルFにファイル格納先ネットワークストレージ識別子PS1を書きこむ。プロセス805では、レプリケーションホストテーブル156には仮想一元化ファイルシステム104が登録されているため、それにCREATE要求を転送することを決定する。プロセス806は、CREATE要求をPS1とVFS1に対してCREATE要求を発行する。

#### 【0117】

PS1におけるCREATE要求処理では、新規ファイルが作成され、応答がレプリケーション手段152の応答収集手段402に返送される。同時にVFS1にもCREATE要求は転送され、応答がレプリケーション手段152の応答収集手段402に返送される。

#### 【0118】

レプリケーション元仮想一元化ファイルシステム103からレプリケーション先仮想一元化ファイルシステム104（VFS1）に転送されたCREATE要

求の処理について説明する。レプリケーション先仮想一元化ファイルシステム 1 0 4 の要求受付手段 1 6 1 は、レプリケーション元仮想一元化ファイルシステム 1 0 3 が発行した C R E A T E 要求を受け付ける。要求受付手段 1 6 1 はレプリケーション手段 1 6 2 を起動する。レプリケーション手段 1 6 2 は、判定手段 1 6 3 を用いて、図 6 に示す判定処理を行ない、プロセス 6 0 2 ではファイルアクセス要求が更新系要求であるかどうかを判定する。C R E A T E 要求は更新系要求であるため、判定結果は“Y E S”となる。次にプロセス 6 0 3 において、マスタフラグが“真”であるかを調査する。仮想一元化ファイルシステム 1 0 4 は、レプリケーション先であるため、マスタフラグは“偽”となっている。そのため、プロセス 6 0 3 の判定結果は“N O”となり、プロセス 6 0 7 におけるレプリケーション判定も“N O”となる。図 5 に示す要求転送処理 5 1 1 では、レプリケーション判定が“N O”であるため、プロセス 5 0 4 を行なう。プロセス 5 0 4 では、ホストテーブル 1 6 5 を用いて、C R E A T E 要求をファイル格納先ネットワークストレージに転送し、プロセス 5 0 6 において、応答をレプリケーション元仮想一元化ファイルシステム 1 0 3 のレプリケーション手段 1 5 2 の応答収集手段 4 0 2 に返送する。

#### 【0 1 1 9】

仮想一元化ファイルシステム 1 0 3 のレプリケーション手段 1 5 2 の応答収集手段 4 0 2 は、V F S 1 と P S 1 からの応答を収集し、一つにして、クライアントに返送してレプリケーション処理を完了する。

#### 【0 1 2 0】

本実施の形態では、C R E A T E 要求および M K D I R 要求をレプリケーションする場合、コピー対象かどうかの判定に、C R E A T E 要求が引数としてもつ初期属性（N F S プロトコルでは、s a t t r に相当）とルールテーブル 4 1 2 を利用する。

#### 【0 1 2 1】

その他の更新系のファイルアクセス要求の場合、一元管理ディレクトリ 1 5 7 のディレクトリおよびダミーファイルの属性情報とルールテーブル 4 1 2 のコピー条件を比較することで、レプリケーションするかどうかを判定する。



### (実施の形態の変形例 1)

上述した代表的実施の形態では、一元管理ディレクトリ 157 とファイルの実体を保持する各ネットワークストレージとの対応付けを行なうために、各ダミーファイルがファイル実体を格納するネットワークストレージの識別子を保持した。ところが、この管理方法では、ファイルアクセスの度にダミーファイルをオープンして、ネットワークストレージの識別子を読み取る必要があるため、ダミーファイルをオープンするオーバーヘッドは大きくなる。

#### 【0122】

そこで、変形例 1 では、一元管理ディレクトリ 157 に配置するダミーファイルにはネットワークストレージ識別子を持たない。そのかわりに、ファイルをグループ化し、ファイルグループ毎にファイルを保持するネットワークストレージの識別子を決定する。

#### 【0123】

図 15 に変形例 1 におけるネットワークストレージシステムの構成を示す。変形例 1 では、仮想一元化ファイルシステム 103 および 104 の仮想一元化装置 110～111 がグループテーブル 1500～1501 を有する。また、一元管理ディレクトリ 1502～1503 のダミーファイルの構造、マッピング手段 1504～1505 の処理が実施の形態とは異なる。さらに、レプリケーション手段 152 で処理するファイルアクセス要求のうち、ネットワークストレージに保持したファイルの実体をアクセスする場合の要求処理が異なる。具体的には、CREATE 要求、READ 要求、WRITE 要求、RENAME 要求、REMOVE 要求、SYMLINK 要求、LINK 要求処理におけるファイル格納先ネットワークストレージを決定するプロセスが異なる。また、SETATTR 要求および GETATTR 要求のうちアクセス対象のファイルオブジェクトがファイルであるものについては、ファイル格納先ネットワークストレージを決定するプロセスが異なる。

#### 【0124】

グループテーブル 1500 の構造を図 16 に示す。グループテーブル 1601 の構造はグループテーブル 1600 と同じである。グループテーブル 1600 は

、ファイルとファイルが属するファイルグループの対応関係を与える。具体的には、ファイル識別子である  $i$  ノード番号  $I$  と、 $i$  ノード番号  $I$  で識別されるファイルグループの識別子  $G$  を与える。1 6 0 1 の行は 4 つのファイルグループの識別番号 0 ～ 3 を示す。1 6 0 2 の行は、1 6 0 1 の各ファイルグループに属するファイルを格納するネットワークストレージ識別子  $PS$  を示す。ファイルグループ 0 に属するファイルの格納先は識別子  $PS 1$  を持つネットワークストレージであり、ファイルグループ 3 に属するファイルの格納先は識別子  $PS 2$  を持つネットワークストレージである。ネットワークストレージの識別子がどのネットワークストレージに対応するかはホストテーブル 1 5 5 を参照すればよい。

#### 【0 1 2 5】

一元管理ディレクトリ 1 5 0 2 には各ネットワークストレージに分散されて保持されるファイルのダミーを置くが、変形例 1 では、ダミーファイルには格納先ネットワークストレージの識別情報を保持しない。一元管理ディレクトリ 1 5 0 3 の構造も一元管理ディレクトリ 1 5 0 2 と同じである。

#### 【0 1 2 6】

マッピング手段 1 5 0 4 は、ファイルとファイルが属するファイルグループの対応を与える。具体的には、一元管理ディレクトリ 1 5 7 で管理されるダミーファイルの識別子である  $i$  ノード番号  $I$  を検索鍵としハッシュ値をファイルグループ識別子とするハッシュ関数で与える。例えば、ファイルを  $N$  個のファイルグループに分割する場合のハッシュ関数として、 $i$  ノード番号  $I$  にファイルグループ数  $N$  のモジュロ演算 ( $N$  で割ったあまり) を適用して得られた値をファイルグループ識別子  $G$  とする。例えば、ファイルグループ数  $N$  が 4 の場合、ファイルグループ識別子は 0 ～ 3 のいずれかをとる。ファイルグループ数  $N$  はネットワークストレージシステムの構成の違いによらず不変である。マッピング手段 1 5 0 5 の構成はマッピング手段 1 5 0 4 と同じである。

次に、変形例 1 におけるレプリケーション処理の流れについて、図 1 5 を用いて説明する。クライアントがファイル更新要求 (WR I T E 要求) をレプリケーション元仮想一元化ファイルシステム 1 0 3 に発行すると、要求受付手段 1 5 1 がその要求を受け付ける。要求受付手段 1 5 1 は、要求を受け付けると、レプリケ

ーション手段152を呼び出し、レプリケーション処理を開始する。レプリケーション処理では、判定手段153によりレプリケーションをするかどうかを判定する。

判定の結果、レプリケーション判定が“YES”である場合には、本発明の実施例と同じように要求転送処理を行なうが、変形例1では、アクセス対象がファイルである場合のファイル格納先のネットワークストレージを、一元管理ディレクトリ1502とグループテーブル1500を利用して特定する点が異なる。

例えば、WRITE要求の場合、図9に示すプロセス903の処理が実施例とは異なる。まず、WRITE要求がアクセス対象とするダミーファイルFのiノード番号を読み出し、マッピング手段1504を用いて、ファイルグループを特定する。次にグループテーブル1500を参照し、ネットワークストレージPSを特定する。

#### (実施の形態の変形例2)

上述の代表的実施の形態では、複数の仮想一元化ファイルシステム間でファイルシステムの部分コピーをリアルタイムに作成するレプリケーション方法について説明した。実施の形態の変形例2では、複数の仮想一元化ファイルシステムではなく、複数ネットワークストレージ間でファイルシステムの部分コピーをリアルタイムに作成する点が実施例とは異なる。

#### 【0127】

図17は、実施の形態の変形例2におけるネットワークストレージシステム全体の構成例を示す。ネットワークストレージシステムは、クライアント1700～1701、ネットワーク1702、レプリケーション装置1710、内部ネットワーク1715、ネットワークストレージ1720～1722で構成する。

#### 【0128】

レプリケーション装置1710以外の構成は本発明の実施の形態と同じであるため、説明を省略する。また、各ネットワークストレージの容量は同じである必要はない。レプリケーション先ネットワークストレージは、コピー対象のディレクトリおよびファイルを保持するための容量を持てばよい。さらに、変形例2では、レプリケーション元もレプリケーション先もネットワークストレージである

が、レプリケーション元はネットワークストレージで、レプリケーション先は仮想一元化ファイルシステムでもよい。

#### 【0129】

レプリケーション装置 1710 は、管理手段 1750、要求受付手段 1751、レプリケーション手段 1752、判定手段 1753、ホストテーブル 1754、レプリケーションホストテーブル 1755、管理テーブル 1756 で構成する。実施の形態の変形例 2 では、レプリケーション装置 1710 に特徴がある。変形例 2 では、レプリケーション装置 1710 は、リモート制御手段およびストレージ装置を持たないが、レプリケーション装置 1710 がリモート制御手段やストレージ装置をもってもよい。また、レプリケーション装置は複数存在してもよい。

#### 【0130】

管理手段 1750 および要求受付手段 1751 の構成および機能は、本発明の実施の形態で説明した管理手段 150 および要求受付手段 151 と同じであるため、説明を省略する。また、変形例 2 では、本発明の実施の形態で説明したマッピング手段 154 は必要ない。

#### 【0131】

ホストテーブル 1754 は、レプリケーション元ネットワークストレージの IP アドレスや名前情報を管理する。レプリケーションホストテーブル 1754 は、レプリケーション先ネットワークストレージを管理する。例えば、変形例 2 では、ホストテーブル 1754 には、ネットワークストレージ 1720 を登録する。また、レプリケーションホストテーブル 1754 には、ネットワークストレージ 1721～1722 を登録する。

#### 【0132】

管理テーブル 1756 は、ホストテーブル 1754 に登録したネットワークストレージのファイルシステムのファイルディレクトリ構造とファイルおよびディレクトリの属性情報を管理する。管理方法は、発明の実施の形態で説明したように、ファイルシステムを利用してもよいし、テーブルを新たに設けてもよい。

#### 【0133】

レプリケーション手段 1752 は、本発明の実施の形態で説明した図 1 に示すレプリケーション手段 152 の構成と同じである。ただし、変形例 2 では、ホストテーブル 1754 の登録内容を調査すれば、どのネットワークストレージがマスタであるファイルおよびディレクトリを保持するかを判別可能であるため、マスタフラグ 411 は必要ない。但し、レプリケーション装置が各ネットワークストレージに含まれる構成の場合、コピー元となるネットワークストレージのレプリケーション装置において、マスタフラグ 411 を“真”に設定する。

#### 【0134】

判定手段 1753 は、ファイルアクセス要求の種類、同期フラグ 410、ルールテーブル 412 に設定した条件を満たすかによってレプリケーション処理をするかどうかを判定する。本発明の実施の形態では、マスタフラグ 411 も利用して条件判定を行っていたが、変形例 2 では、マスタフラグ 411 を利用しない。

#### 【0135】

判定手段 1753 に同期フラグ 410 およびルールテーブル 412 を備えること、及びこれら同期フラグ、ルールテーブルを用いた判定の手法は上述の代表的実施の形態と同じである。

#### 【0136】

変形例 2 におけるレプリケーション処理では、まず判定手段 1753 を用いて判定処理を行なう。

#### 【0137】

判定処理では、まずファイルアクセス要求が更新系要求かどうかを判定し、次に、同期フラグが“真”かどうかを判定し、ルールテーブルによりアクセス要求の対象がコピー対象であるかどうかを判定する。いずれも“YES”の場合にレプリケーション判定が“YES”となる。

#### 【0138】

レプリケーション判定が“YES”である場合には、要求転送処理において、ファイルアクセス要求をホストテーブル 1754 に登録したネットワークストレージと、レプリケーションホストテーブル 1755 に登録したネットワークスト

レージに転送する。レプリケーション判定が“NO”である場合には、ホストテーブル 1 7 5 4 に登録したネットワークストレージにのみファイルアクセス要求を転送する。

#### (実施の形態の変形例 3)

上述の代表的実施の形態では、仮想一元化ファイルシステムに保持したデータの部分コピーをリアルタイムに作成するレプリケーション方法について説明した。変形例 3 では、ファイルのレプリケーションを行なっている複数の仮想一元化ファイルシステムにおいて、クライアントが管理手段 1 5 0 を介してレプリケーション元仮想一元化ファイルシステムに設定する設定情報のコピーも作成する。コピー対象の設定情報としては、アカウントやセキュリティ情報や時刻情報がある。あるいは、IP アドレスや名前情報を含むこともある。設定情報は、更新される度に特定ファイルにパッキングする。特定ファイルはルールテーブル 4 1 2 においてコピー対象として設定することで、設定が変更された場合には必ずコピーされるようにする。仮に、レプリケーション元仮想一元化ファイルシステム 1 0 3 に障害が発生し使用できなくなった場合には、管理者がレプリケーション先仮想一元化ファイルシステム 1 0 4 をマスタとして設定するのに応答して、管理手段 1 6 0 は、特定ファイルをアンパッキングし、設定情報を設定する。

#### 【0 1 3 9】

変形例 3 では、設定情報をレプリケーション先仮想一元化ファイルシステム 1 0 4 に特定ファイルとしてコピーする方法を記載したが、クライアント計算機に設定情報のバックアップを取得しておく方法でもよい。あるいはクライアントが明示的に設定情報をレプリケーション元とレプリケーション先の仮想一元化ファイルシステムに入力してもよい。

#### 【0 1 4 0】

##### 【発明の効果】

以上のように、複数のネットワークストレージで構成される分散システムにおいて、要求受付手段が一括してクライアントからのファイルアクセス要求を受け付け、管理テーブルを利用してレプリケーションするかどうかを判定し、コピー対象のファイルに対するファイルアクセス要求のみを、レプリケーション元ネッ

トワークストレージと、レプリケーション先ネットワークストレージの両方に転送することで、レプリケーション元ネットワークストレージのファイルシステムの部分コピーをリアルタイムに別のネットワークストレージに作成することができる。これにより、大容量ネットワークストレージに保持したデータのバックアップを効率よく行なうことができる。

#### 【0141】

また、複数ネットワークストレージと仮想一元化装置で構成される仮想一元化ファイルシステムにおいて、クライアントからファイルアクセス要求が発行されると、仮想一元化装置がそのファイルアクセス要求を受け付け、一元管理ディレクトリを利用してアクセス対象ファイルの存在するネットワークストレージを特定し、かつファイルアクセス要求のアクセス対象であるファイルがコピー対象であるかどうかを判定し、コピー対象のファイルに対するファイルアクセス要求である場合にのみ、アクセス対象のネットワークストレージとレプリケーション先仮想一元化ファイルシステムの仮想一元化装置に対してファイルアクセス要求を同時に転送することで、レプリケーション対象元仮想一元化ファイルシステムの部分コピーをリアルタイムに別の仮想一元化ファイルシステムに作成することができる。これにより大容量仮想一元化ファイルシステムのデータバックアップを効率良く行うことができる。

#### 【0142】

さらに、レプリケーション対象仮想一元化ファイルシステムを構成する仮想一元化装置が、コピー対象のファイルおよびディレクトリと、レプリケーション先仮想一元化ファイルシステムに格納したファイルおよびディレクトリとの整合性をとる手段を持つため、障害発生によりレプリケーション元とレプリケーション先仮想一元化ファイルシステム間でコピー対象のファイルおよびディレクトリの整合性が崩れても、障害復旧時に簡単に再度整合性をとることができる。これにより、障害復旧に必要な管理コストを削減できる。

#### 【図面の簡単な説明】

#### 【図1】

本発明の実施の形態であるネットワークストレージシステムの全体構成を示すブロック図である。

【図 2】

上記実施の形態におけるホストテーブル 1 5 5 の構成を示す図である。

【図 3】

上記実施の形態におけるレプリケーションホストテーブル 1 5 6 の構成を示す図である。

【図 4】

上記実施の形態におけるレプリケーション手段 1 5 2 の構成を示す機能ブロック図である。

【図 5】

上記実施の形態におけるレプリケーション処理を示すフロー図である。

【図 6】

上記実施の形態における判定処理 5 1 0 を示すフロー図である。

【図 7】

上記実施の形態におけるファイルハンドル 7 0 0 の構成を示す図である。

【図 8】

上記実施の形態における C R E A T E 要求処理を示すフロー図である。

【図 9】

上記実施の形態における W R I T E 要求処理を示すフロー図である。

【図 1 0】

上記実施の形態における M K D I R 要求処理を示すフロー図である。

【図 1 1】

上記実施の形態における L O O K U P 要求処理を示すフロー図である。

【図 1 2】

上記実施の形態における R E A D 要求処理を示すフロー図である。

【図 1 3】

上記実施の形態における整合処理を示す図である。

【図 1 4】



上記実施の形態におけるレプリケーション設定手順を示すフロー図である。

【図 1 5】

実施の形態の変形例 1 であるネットワークストレージシステムの全体構成を示すブロック図である。

【図 1 6】

上記変形例 1 のグループテーブル 1 5 0 0 の構成を示す図である。

【図 1 7】

実施の形態の変形例 2 であるネットワークストレージシステムの全体構成を示すブロック図である。

【符号の説明】

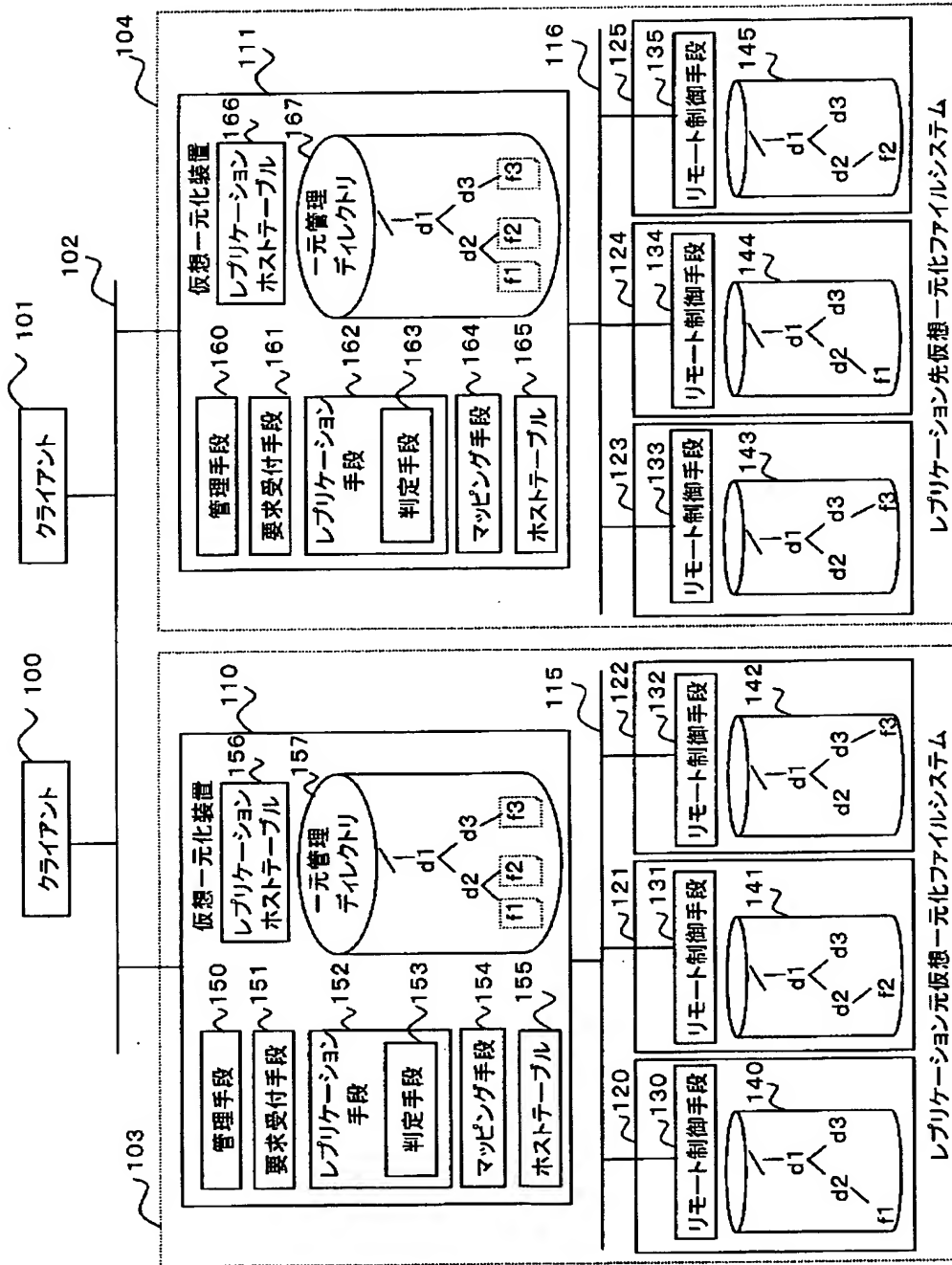
1 0 0 ~ 1 0 1 …クライアント、1 0 2 …ネットワーク、1 0 3 …レプリケーション元仮想一元化ファイルシステム、1 0 4 …レプリケーション先仮想一元化ファイルシステム、1 1 0 ~ 1 1 1 …仮想一元化装置、1 1 5 ~ 1 1 6 …内部ネットワーク、1 2 0 ~ 1 2 5 …ネットワークストレージ、1 3 0 ~ 1 3 5 …リモート制御手段、1 4 0 ~ 1 4 6 …ストレージ装置、1 5 0 …管理手段、1 5 1 …要求受付手段、1 5 2 …レプリケーション手段、1 5 3 …判定手段、1 5 4 …マッピング手段、1 5 5 …ホストテーブル、1 5 6 …レプリケーションホストテーブル、1 5 7 …一元管理ディレクトリ、1 6 0 …管理手段、1 6 1 …要求受付手段、1 6 2 …レプリケーション手段、1 6 3 …判定手段、1 6 4 …マッピング手段、1 6 5 …ホストテーブル、1 6 6 …レプリケーションホストテーブル、1 6 7 …一元管理ディレクトリ、4 0 1 …設定手段、4 0 2 …要求処理手段、4 0 3 …応答収集手段、4 0 4 …整合手段、4 1 0 …同期フラグ、4 1 1 …マスタフラグ、4 1 2 …ルールテーブル、1 5 0 0 ~ 1 5 0 1 …グループテーブル、1 7 5 6 …管理テーブル。

【書類名】

図面

【図 1】

図 1



【図 2】

図2

ホストテーブル155

NS120	NS121	NS122	201
/	/	/	202
PS1	PS2	PS3	203

【図 3】

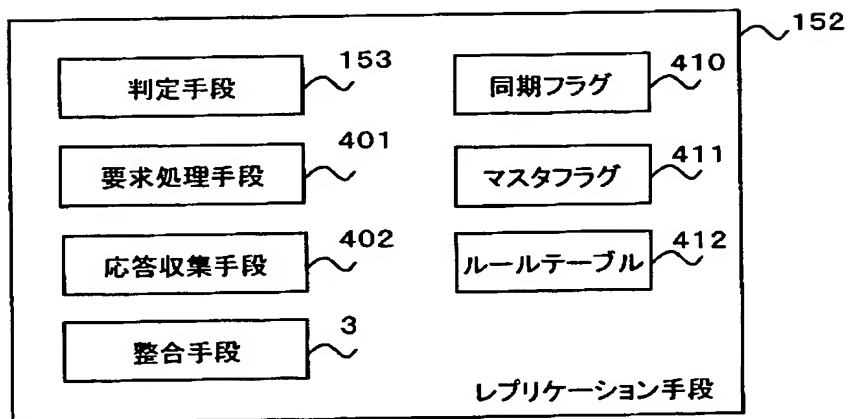
図3

レプリケーションホストテーブル156

VFS104	● ● ●	XXXX	301
/			302
VFS1	● ● ●	VFSk	303

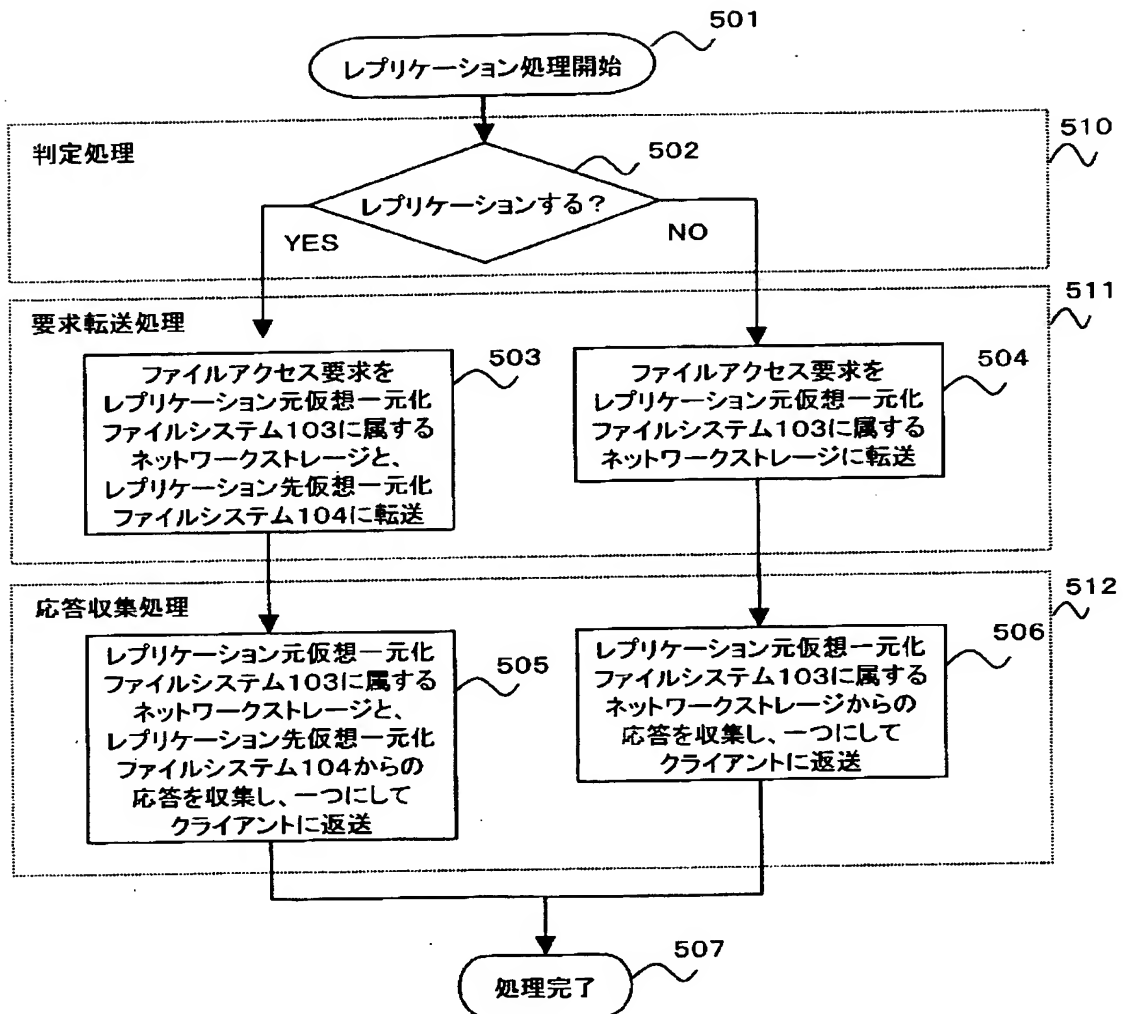
【図 4】

図4



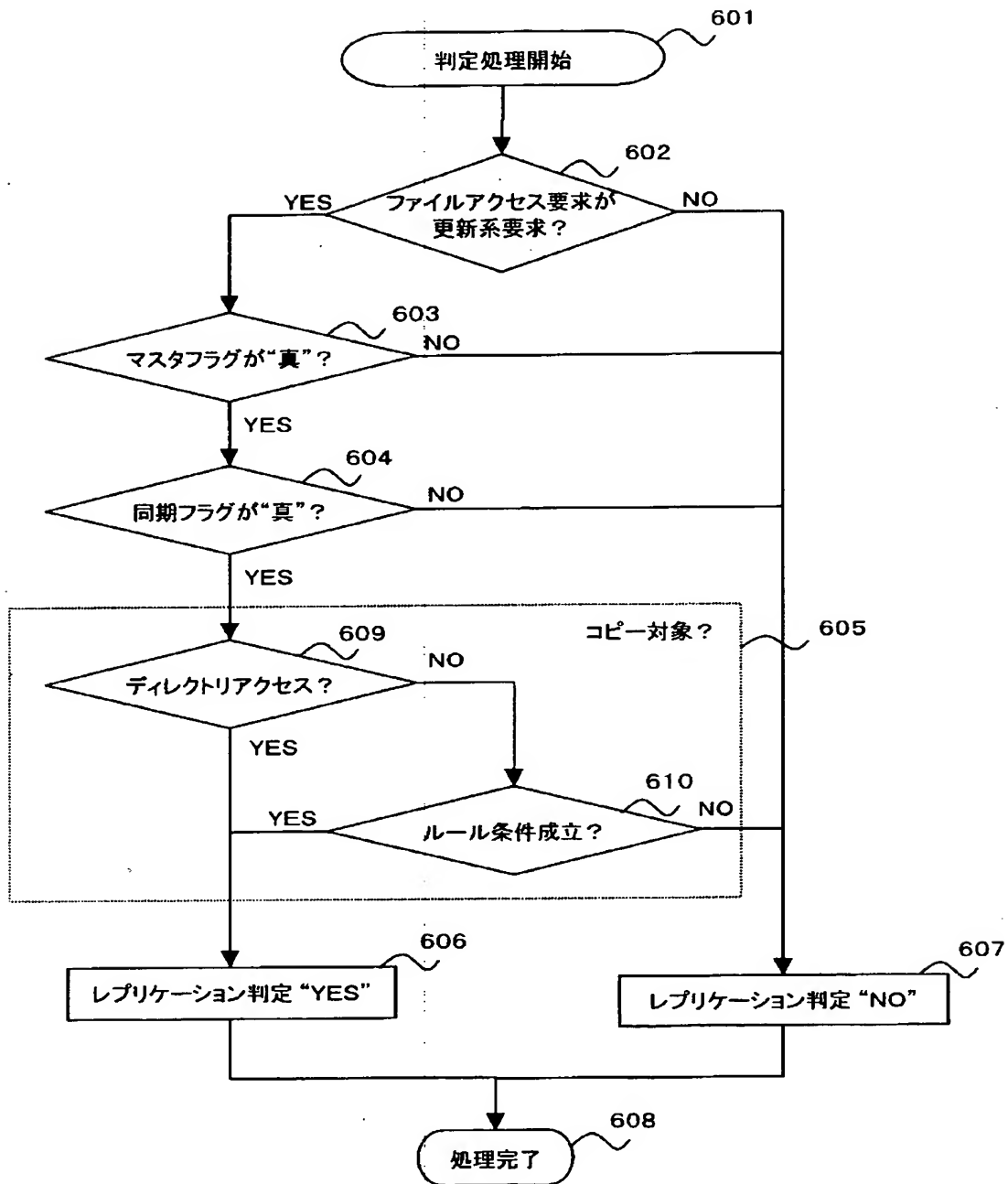
【図 5】

図5



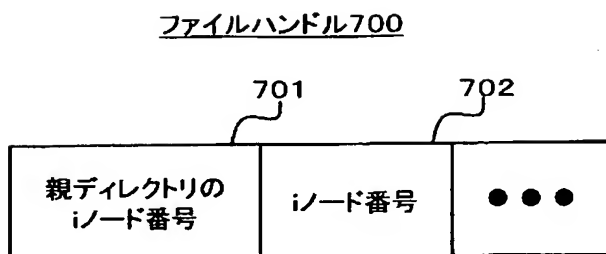
【図 6】

図6



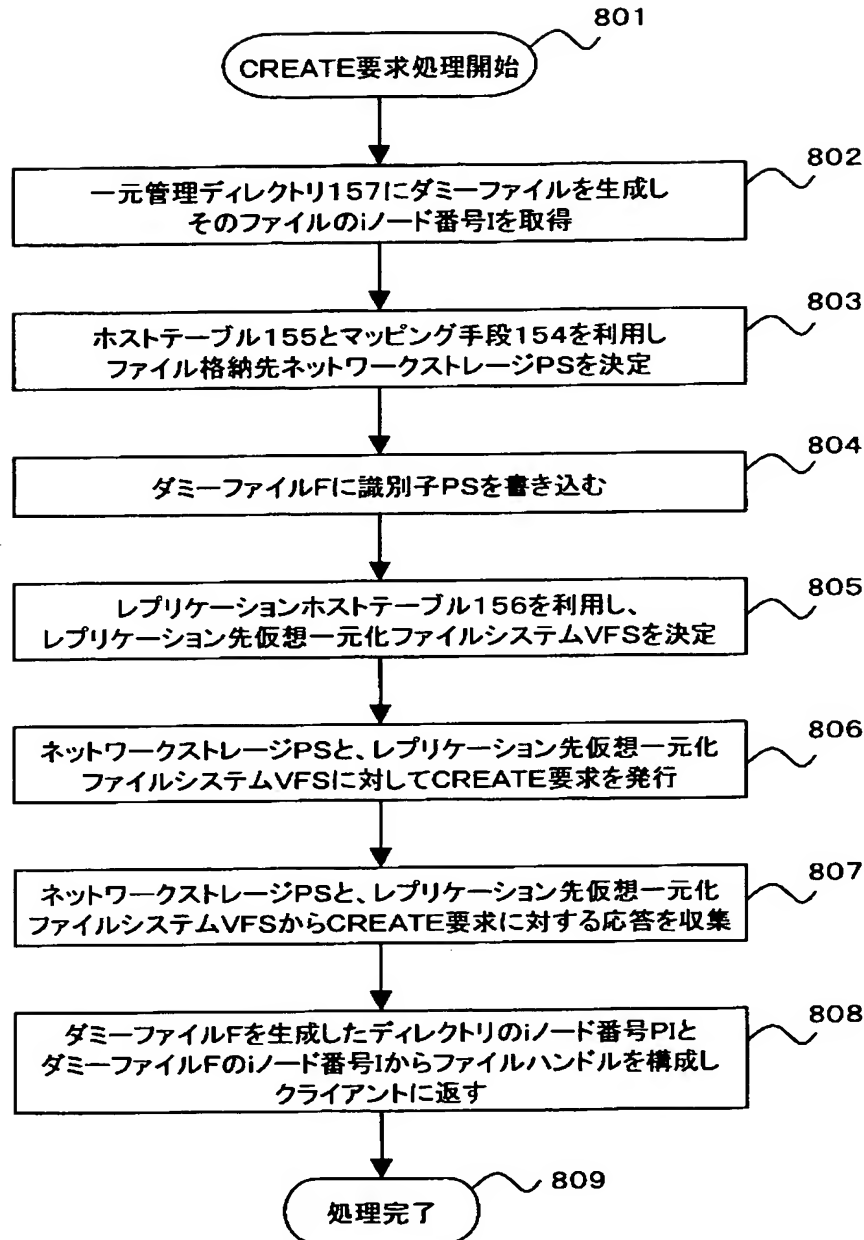
【図 7】

図7

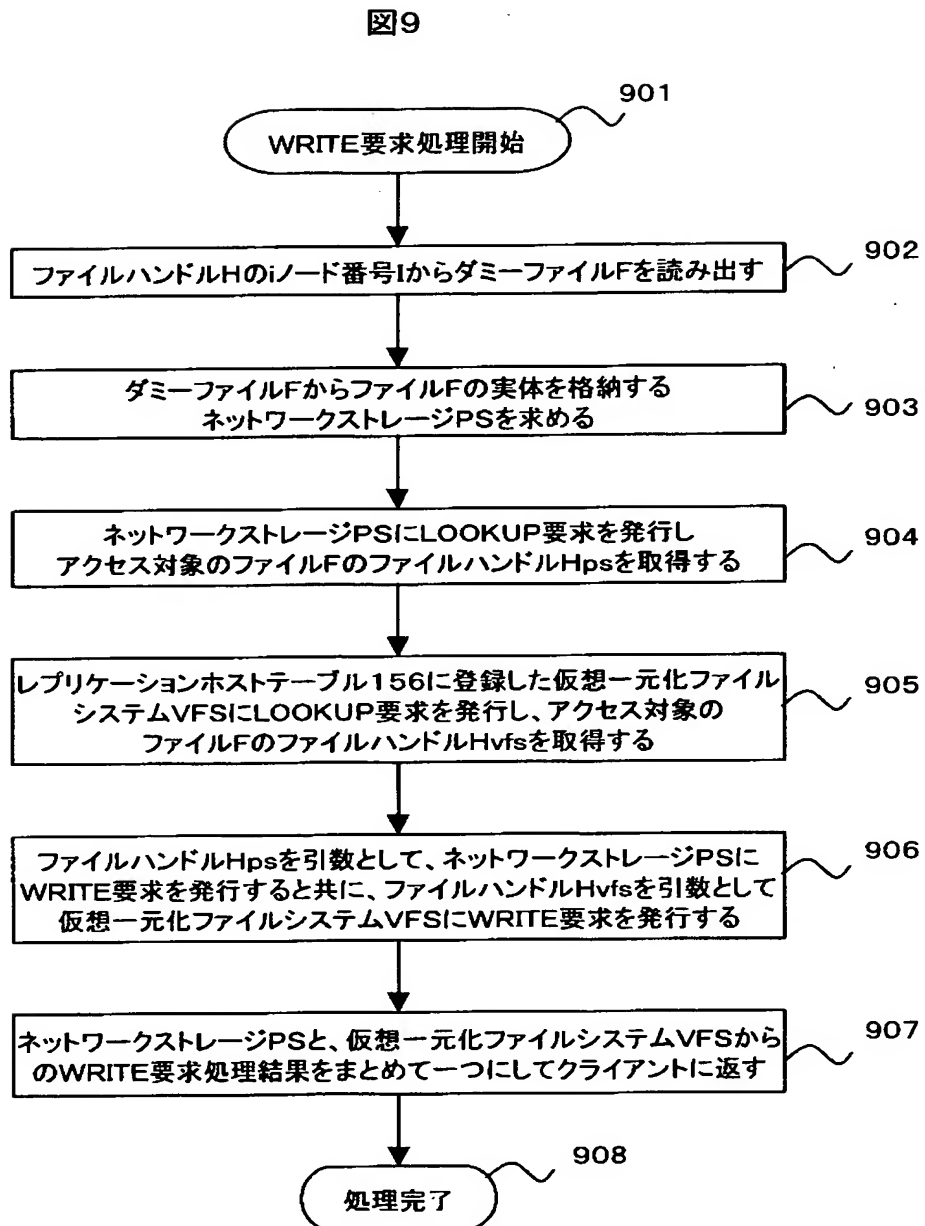


【図 8】

図 8



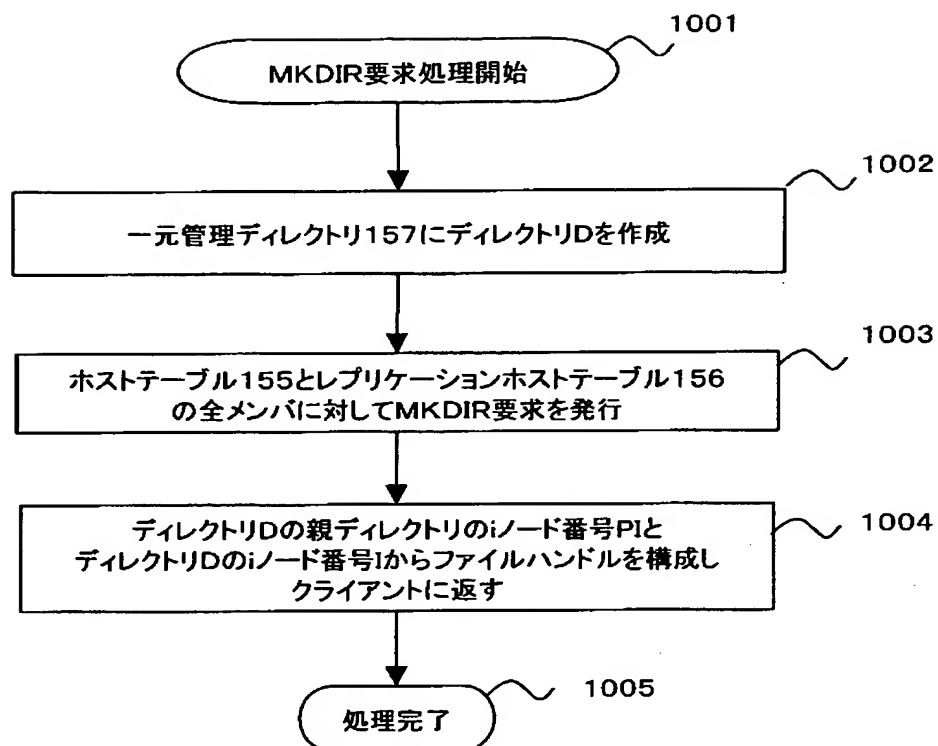
【図 9】



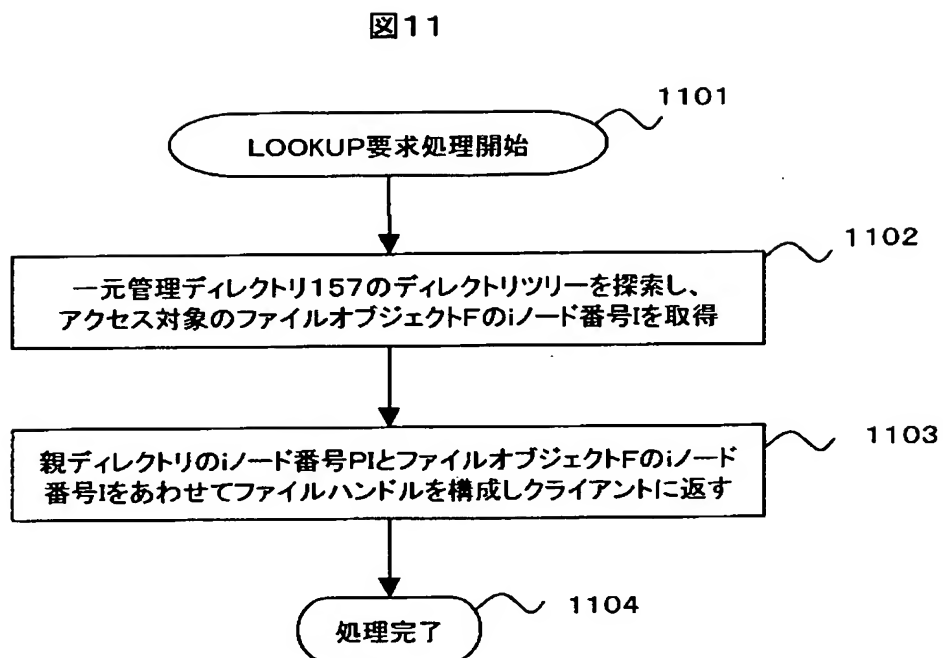


【図10】

図10

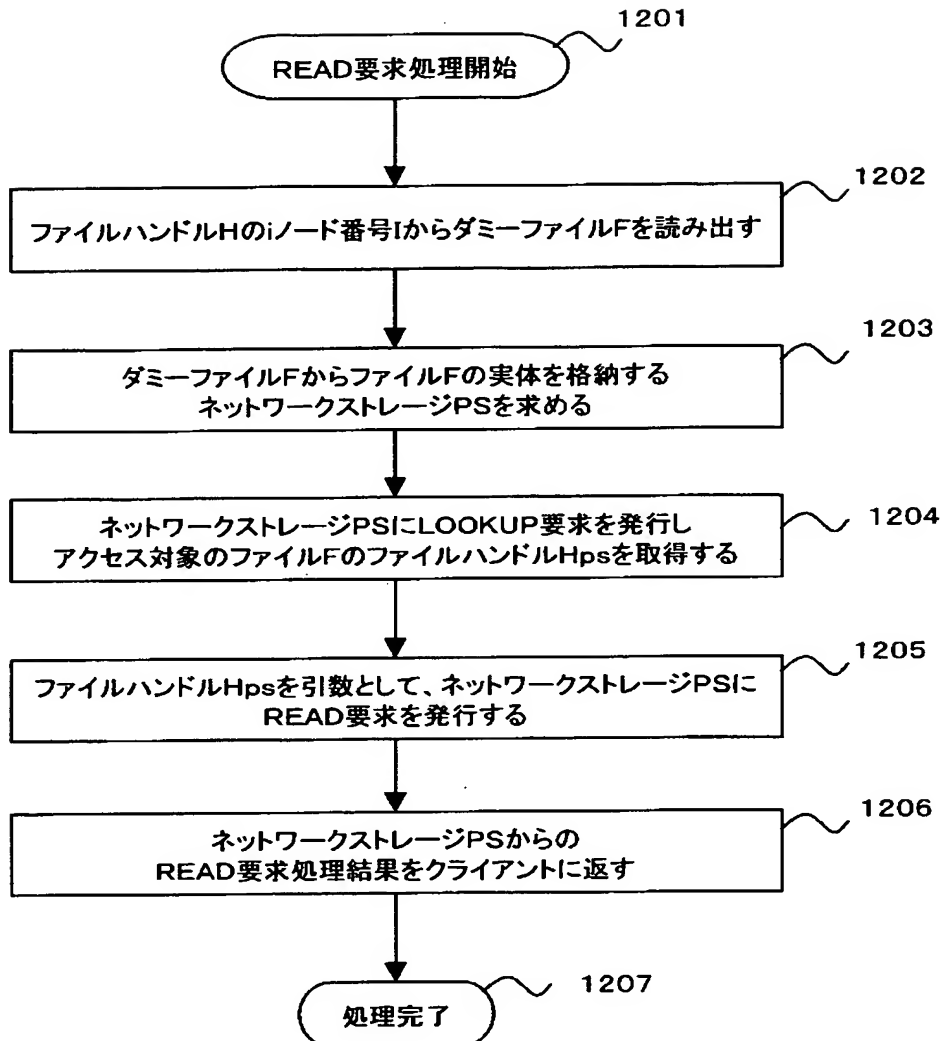


【図 11】

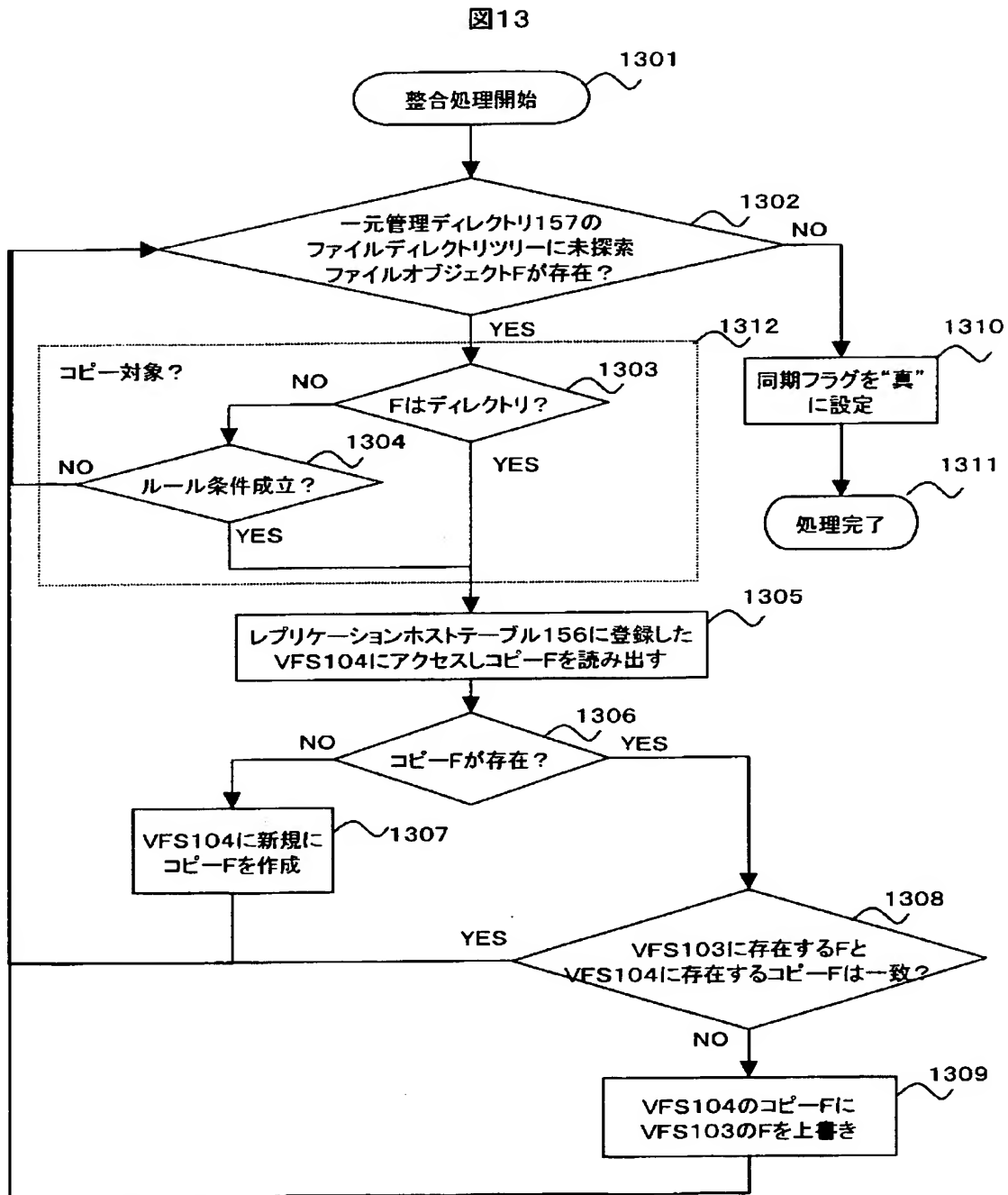


【図 12】

図12

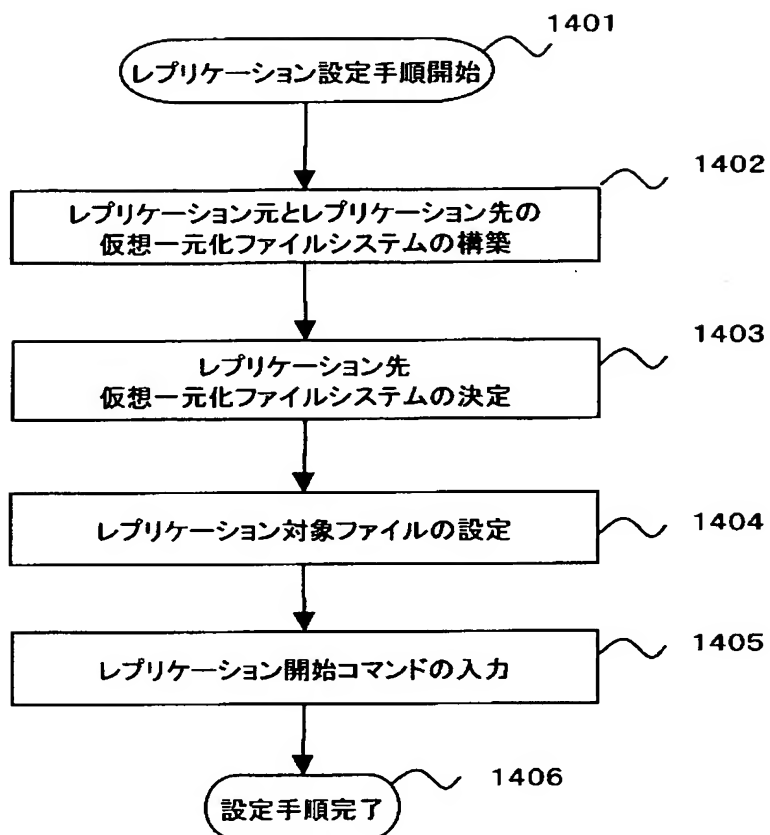


【図13】

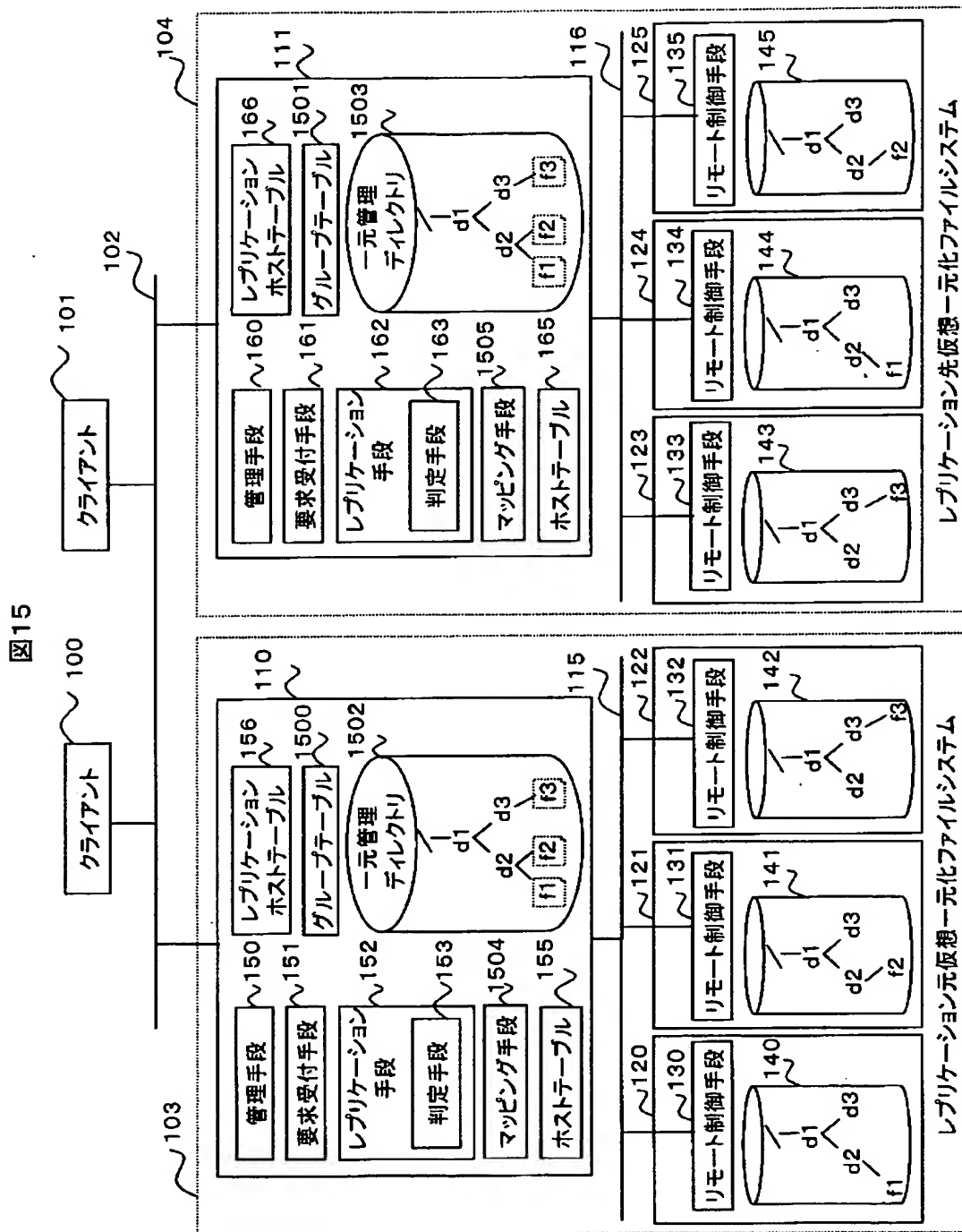


【図 14】

図 14



【図 15】



【図16】

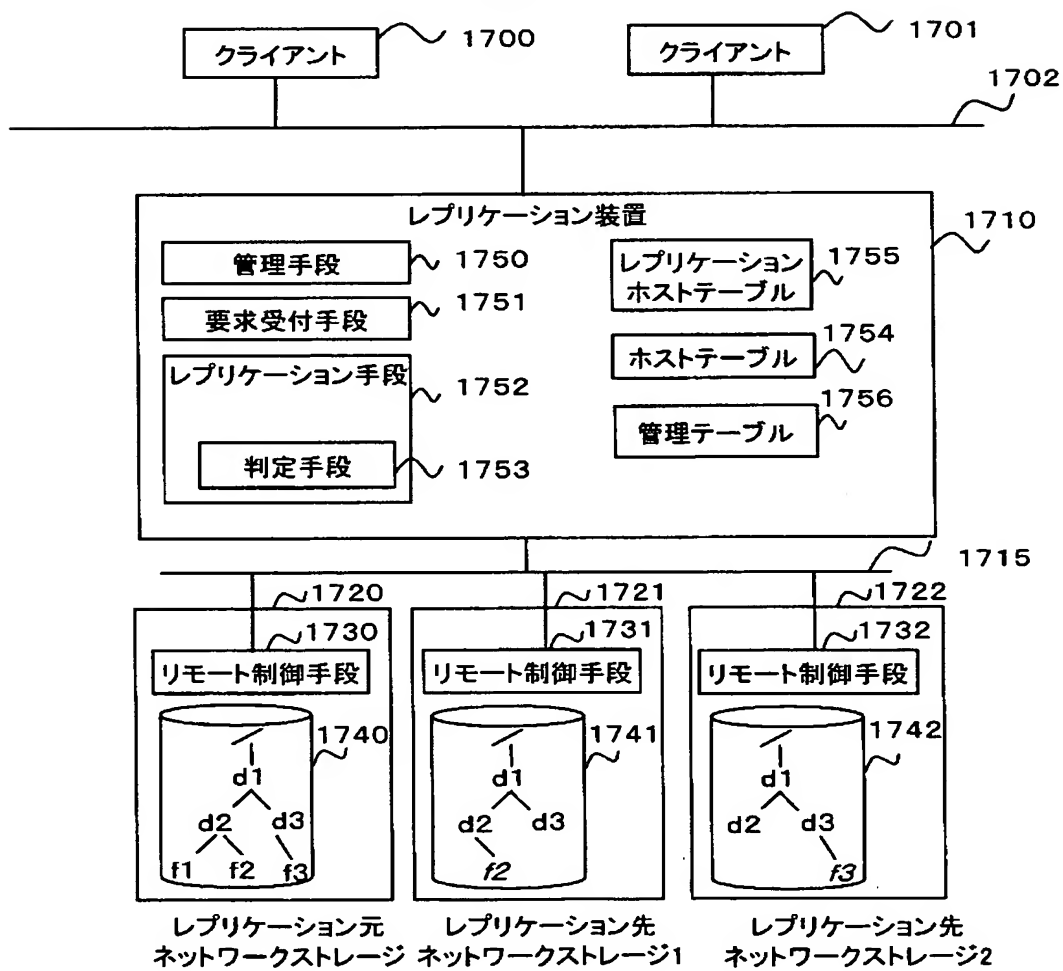
図16

グループテーブル1500

0	1	2	3	1601
PS1	PS1	PS2	PS2	1602

【図17】

図17







【書類名】 要約書

【要約】

【課題】 複数の仮想一元化ファイルシステム間でレプリケーションの対象となる仮想一元化ファイルシステムの部分コピーをリアルタイムに作成するファイルレプリケーション方法を提供する。

【解決手段】 レプリケーション元仮想一元化ファイルシステムに、ファイルディレクトリの構造と属性情報を管理する一元管理ディレクトリ、コピー対象のファイルを設定するレプリケーション情報、クライアントからの要求を一元的に受け付ける要求受付手段、一元管理ディレクトリとレプリケーション情報に従ってレプリケーションするかどうかを判定し、判定結果によりファイルアクセス要求をレプリケーション元とレプリケーション先の仮想一元化ファイルシステムに同時に転送するレプリケーション手段を設ける。

【効果】 複数の仮想一元化ファイルシステムにおいて、レプリケーション元ファイルシステムの部分コピーをレプリケーション先にリアルタイムに作成できるため、大容量システムのデータバックアップを効率良く行うことができる。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願 2 0 0 3 - 0 2 1 3 0 1
受付番号	5 0 3 0 0 1 4 3 7 4 0
書類名	特許願
担当官	土井 恵子 4 2 6 4
作成日	平成 1 5 年 1 月 3 1 日

< 認定情報・付加情報 >

【提出日】	平成 15 年 1 月 30 日
-------	------------------

次頁無

特願 2 0 0 3 - 0 2 1 3 0 1

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 5 1 0 8 ]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所